

STIC-Biotech/Ch mLib

79027

Fr m: Mehta, Ashwin
Sent: Wednesday, October 30, 2002 9:06 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

STIC,

Please search the commercial and interference databases for the following from 09/896,186:

- 1) the nucleotide sequence of SEQ ID NO: 23
- 2) the amino acid sequence of SEQ ID NO: 24

My mail room is 9E12, office 9E07, art unit 1638.

Thank you,
Ashwin

Ashwin Mehta
United States Patent and Trademark Office
Biotechnology Patent Examiner
703-306-4540

RECEIVED
OCT 30 2002
STIC

Point of Contact
P. Sheppard

Searcher's telephone number: (703) 308-4499

Phone: _____

Location: _____

Date Picked Up: _____

Date Completed: 11/7/02

Searcher Prep/Review: _____

Clerical: _____

Online time: _____

TYPE OF SEARCH:

NA Sequences: _____

AA Sequences: _____

Structures: _____

Bibliographic: _____

Litigation: _____

Full text: _____

Patent Family: _____

Other: _____

VENDOR/COST (where applic.)

STN: _____

DIALOG: _____

Questel/Orbit: _____

DRLink: _____

Lexis/Nexis: _____

Sequence Sys.: _____

WWW/Internet: _____

Other (specify): _____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 14:04:54 ; Search time 2044 Seconds

(without alignments)
8311.676 Million cell updates/sec

Title: US-09-896-186B-23

Perfect score: 1049

Sequence: 1 accaagacatatttatt.....ttaaccgtccagaactag 1049

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estfpl:*
7: em_esttro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_fod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	288.8	27.5	435	9 AU226180	AU226180 AU226180
C 2	273.8	26.1	840	10 BE659005	BE659005 GA00008A
C 3	261	24.9	704	12 BG441595	BG441595 GA_Ea001
C 4	245	23.4	552	10 AM201789	AM201789 sf07c10.y
C 5	243.8	23.2	671	9 AU237507	AU237507 AU237507
C 6	241.2	23.0	467	10 AM832139	AM832139 sm20b11.y

7	235.4	22.4	661	14	BO799068	EST_1237
8	233.8	22.3	535	13	BM406124	EST580451
9	223	21.3	506	12	BG133320	EST466112
10	213.6	20.4	531	13	BM094501	sa116903.
11	172.2	16.4	515	13	BM358892	GA_Ea001
12	169.6	16.2	517	13	BM358891	GA_Ea001
13	167.4	16.0	555	14	BM731125	sa168a05.
14	162	15.4	595	13	Bu213300	Bu713300
15	156.8	14.9	499	13	BI430308	EST535799
16	156.6	14.9	556	14	BM890816	BM890816 sam08c02.
17	156	14.9	563	13	BM527093	sa149e10.
18	152	14.5	531	10	AM460204	sh56c10.y
19	151	14.4	684	12	BE589996	BE589996 EST497838
20	143.2	13.7	625	12	BE999775	EST431498
21	142.4	13.6	560	12	BE157629	sa088c03.
22	140.8	13.4	572	9	AU235472	AU235472 AU235472
23	134.4	12.8	680	13	Bu321563	Bu321563 Bu321563
24	129.6	12.4	407	12	BE821184	BE821184 GM700024A
25	129	12.3	220	17	BH849559	BH849559 SALK_0698
26	124.6	11.9	431	9	A1748378	A1748378 sb52a06.y
27	117	11.2	445	10	AM119478	AM119478 sd46a06.y
28	100.6	9.6	494	10	AM720170	AM720170 LjNEST16b
29	99	9.4	654	13	Bu220795	Bu220795 Bu220795
30	97.8	9.3	660	17	BE1114	BE1114 T20K8PF TAM
31	96	9.2	600	13	Bu316072	Bu316072 Bu316072
32	84.6	8.1	550	14	BM785803	BM785803 K-EST0064
33	76.2	7.3	405	17	CNS0008H	AL09167 Arabidops
34	66.8	6.4	533	17	AO580338	AO580338 T135971b
35	65.4	6.2	473	17	AO866027	AO866027 nbe00026P
36	65.4	6.2	901	17	AO867999	nbe00018H
37	62	5.9	342	17	AO841745	AO841745 T132158b
38	62	5.9	457	17	BH168898	BH168898 SALK_0003
39	57.2	5.5	447	13	BI420478	BI420478 LjNEST57f
40	56.6	5.4	586	13	BM191106	BM191106 da32b10.
41	55.6	5.3	569	17	AO580359	AO580359 T136000b
42	50.2	4.8	151	10	BE022823	sm89c11.y
43	47	4.5	711	13	Bu169253	Bu169253 Bu169253
44	46.6	4.4	620	9	AL647838	AL647838
45	45.8	4.4	161	17	B24158	F19B10TR IG

ALIGNMENTS

RESULT 1
AU226180/c 435 bp mRNA linear EST_23-APR-2002
LOCUS AU226180
DEFINITION AU226180 RAF14 Arabidopsis thaliana cDNA clone RAF14-14-A20 3 ,
ACCESSION AU226180
VERSION AU226180.1 GI:19740827
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 435)
REFERENCE
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J.,
Itoh, M., Ishii, Y., Atakawa, T., Shibata, K., Shinagawa, A., Muramatsu,
'M., Hayashizaki, Y. and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rkc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI

and XhoI was ligated to modified Lambda FLG-1 vector (Carlini et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified Bluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

```

1.435
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAF14-14-820"
/clone_1lb="RAF14"
/tissue_type="root"
/lab_host="DH10B"
/notes="Site_1: BamHI; Site_2: SalI"

```

BASE COUNT 117 a 85 c 85 g 148 t

ORIGIN

Query Match

Best Local Similarity 90.3%; Pred. No. 2,2e-61; Matches 327; Conservative 0; Mismatches 17; Indels 18; Gaps 1;

```

QY 706 GAGATAAAAATGGGCGCTTGCCTCACTAAGACACTGTTTGCAGAGAGCTCCTGA 765
    |||||||
Db 435 GAGATAAAAATGGGCGCTTGCCTCACTAAGACACTGTTTGCAGAGAGCTCCTGA 376
QY 766 AGCCAAACAGAAATCAGGCTTGGGAACTGGAGTTTATCCTCTGTCAAGACAGCTTAC 825
    |||||||
Db 375 AGCCAAACAGAAATCAGGCTTGGGAACTGGAGTTTATCCTCTGTCAAGACAGCTTAC 316
QY 826 AATACGACAGCAAGGATGCTTATGCTTCAATGATGCTTCAAGGTTCTTAAGACCTTC 885
    |||||||
Db 315 AATACGACAGCAAGGATGCTTATGCTTCAATGATGCTTCAAGGTTCTTAAGACCTTC 256
QY 886 CTGATGCTGCTGCTGCTCACTAAGTGAAGAGAGAGCTTAAAGTTAGGCTTAAACCCC 945
    |||||||
Db 255 CTGATGCTGCTGCTGCTCACTAAGTGAAGAGAGAGCTTAAAGTTAGGCTTAAACCCC 196
QY 946 AAGATTTAGCATCAAAATGATATGATACACCTAATCTAGTCAAGTGAATGCAATTCCT 1002
    |||||||
Db 195 AAGATTTAGCATCAAAATGATATGATACACCTAATCTAGTCAAGTGAATGCAATTCCT 136
QY 1003 -----GTGATATTTGATCTAGTCTGCTGCTTAAAGCTTAAAGCTTAAAGCTTAA 1047
    |||||
Db 135 GGAATTACAAAGATATTGAACACAGATATTAGTTGGCGTCCATTAATCTTCCAGAAACT 76
QY 1048 AG 1049
    ||
Db 75 AG 74

```

RESULT 2
BE659005/c 840 bp mRNA linear EST 24-MAY-2001
LOCUS GM700008A10G11 Gm-r1070 Glycine max cDNA clone Gm-r1070-2997 3',
DEFINITION mRNA sequence.

ACCESSION BE659005
VERSION BE659005.1 GI:9984897
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 840)
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V., Erpelting, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H., A Functional Genomics Program for Soybean (NSF 9872565)
TITLE Unpublished (1999)
JOURNAL Other ESTs: AW460204 corresponding to Gm-cl015-3787 (5')
COMMENT Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genome systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.

FEATURES

source

```

1.840
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="Gm-r1070-2997"
/clone_1lb="Gm-r1070"

```

/note="The library Gm-r1070 is a sequence-driven, rerecked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were rerecked to form library Gm-r1070. The cDNA clones of the rerecked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the Laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://www.cbc.umn.edu/ResearchProjects/soybean/index.html>. Rerecking was performed by Genome Systems, St. Louis, <http://www.genomesystems.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.life.uiuc.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'.

BASE COUNT 242 a 167 c 158 g 240 t 33 others
ORIGIN

Query Match 26.1%; Score 273.8; DB 10; Length 840;
Best Local Similarity 66.1%; Pred. No. 1.3e-57;
Matches 392; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

QY 300 AGGAATTTCCACCAATGAGCTTGGTGGTACGATTTGTATAGCAAGACTGCTACTGAG 359
    |||||
Db 727 ATGACATTTGCTGTATGGAAGTTAGTGTCACAAATTTCTTATGACGAGCTTTGATGCT 668
QY 360 GTTGATTAAGGAGCAATGCAAGCTTAAAGTTCTTGATACCAAGAGAGATGATCGGA 419
    |||||
Db 667 GTAGAGAAAGCTGCACAAAGCNCCTTACAAATCTCCAGAAAGAAAGCAGCAGATGATG 608
QY 420 ATAGCTTTTGGCTTGGATATGATATGATGAGCAAGTTTGAAGAGGTTGTCGCCG 479
    |||||
Db 607 CAACATGCAATTTGATTTGACATTTGATGAGTGAAGCCCNMCNTGAGAAAGGTGTTCTCC 548
QY 480 GGGAGAGTTGCGCATCTCCAGATATGATAGATTAATTAATTTGATGATTTATGATATT 539
    |||||
Db 547 GGAAGGTAGCACTGATGACATATGTGTGACACATGACATTTGATTTCTACATCTTA 488
QY 540 TTTCAATTTGATTTCCCTCAAGACTCCAAACATTTATTAAGATTTCAACACTTGAAG 599
    |||||
Db 487 ATTCATTTGGAATCCCTCAAAATTTACAGCTTTCTTGAAGATCCACAGCTTGAG 428
QY 600 GTAGATTTGGAATTTAGTGCTGTGAGAGCTTTTCCATGACATAGAGATTAATGATATC 659
    |||||
Db 427 GTTGGAGCTGGGATTTAGTGCTGTGAGAGCTTTTGAAGATTAATTAATATCTGTT 368

```


OY	660	AAAGATGTTGAGATCTTTACAGATTTAGCCACCCAAAAAATTGGTGGAGATATAAAAAATGG	719
Db	367	AAAGGTGTGACGGATCTTTCTTTATCATCAATCAAGCTTGGTGGAGATCATAAAGTTG	308
OY	720	GCCCTTGGCCTCACTAAGCAGACACTTGTGTCGAAGAAGCTCTCAAGCAACAGCAATC	779
Db	307	GCTTTCGATCTTTTGACATCGAAAACTTCTATCAAAACACAGCTTAAAAAGCCCAACAAATA	248
OY	780	AGGCTTGGGAACCTGGAGCTTTATCCTCTGTCAAAAGCAGCACTTACAATATGCGACCAACG	839
Db	247	AGACGTGGGAAATTTGGAGAGGCTCTCTTTTGTCAAAAGGAGCACTAGAGATATCTCGACACA	188
OY	840	GATGCTTATNGCTCATCGATCACTTATACAAAGTTCTTAAGGACCTTCGTATG	892
Db	187	GATGGCTTTTGCTTCTTGCTGTCTTTATATACGAGATTAAGATCTCCCGGACGC	135

RESULT 3	LOCUS	DEFINITION
BG441595	BG441595	704 bp mRNA linear EST 15-MAR-2001 GA_Ea0013022f Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum cDNA clone GA_Ea0013022f, mRNA sequence.

ACCESSION	BG441595
VERSION	BG441595.1
KEYWORDS	GI:13351247
SOURCE	EST.
ORGANISM	Gossypium arboreum.
	Gossypium arboreum

REFERENCE
1 (bases 1 to 704)

AUTHORS	TITLE	JOURNAL
Wang, R. A., Fritsch, D., Yu, Y., Maiti, D., Rambo, T., Simmons, J., Henry, D., Wood, T. C., Leslie, A., and Wilkins, T. A.	An integrated analysis of the genetics, development, and evolution of the cotton fiber	Unpublished (2000)

COMMENT

Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293

FEATURES	Email: twing@clemons.edu
	Seq primer: TAAATACGACATCCACTATAGG
	High quality sequence stop: 701.
source	Location/Qualifiers
	1..704

```

/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0013022f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

```

```

BASE COUNT      188 a      144 c      155 g      217 t
ORIGIN
Query Match      24.9%      Score 261; DB 12;      Length 704;
Best Local Similarity 63.7%      Pred. NO. 1.9e-54;
Matches 434; Conservative 0; Mismatches 233;      Indels 12; Gaps 2;

```

[illegible]

OY	265	CTAATAAAGCATTCCTCCTCCCGCTGCGAGCA-----GGAATTTTCCAGCA	315
Db	144	CTTCTTTTCCGCTTCCTCGCTGCTCTTCATCAATAACAGGCTTGATATCCGCCGT	203
OY	316	TGAGGTTTGGTGAGATTTTGTATACCAAGACTGCTACTGAGGTATATAAGCAGCA	375
Db	204	TGAGATTTGGAGGCTCATATTTTGTATACCTTTACGGAAGATGAGGTACAAAAATGCTGCA	263
OY	376	TGCACCTTATTAAGTCTTGATACCAAGAGAGATGAATCTGGAAATCTTTTGTGGCT	435
Db	284	TGGACCTTATTAATAAATGTTGAAATTAAGAAAGAAATGGGTCAAGTGTCTTAGGGT	323
OY	436	TGGATATTGAGTGGAGACCAAGTTTAAAGAAAGTGTCTCCCGGGAAAGTTGGCACTG	495
Db	324	TTGATATTGAGTGGAGAGCCCTCTTTCCAAAGAAATTTTGCCTGGAGAGCTGGGTGA	383
OY	496	TTCACATATTGCTTCAATATGTAATTTATTTGATGATATCATATTTTATATTCGTGATCC	555
Db	384	TGCAGATATTGTTGTACAGCTCAGTATTGTTATGTATATGCAATATTTTATATTCGGCATAC	443
OY	556	CTCAAAAGTCTCCAAACATCTTATTTGAAGATTCAACCTGTAAGGTAGATAATTGGAATTG	615
Db	444	CTCAAAAGTCTGAGGTCCTCTTTGAGGACTCCGAATATTATAAAGTTGAGGTGCAATTG	503
OY	616	ATGTTGACTCTGTGAAGCTTTCCATGACTATGAGGTTAGTATCAAGAATGTTGAGATC	675
Db	504	ATGGGAGTGTCTCAGAGGTGTTCAGCTGACTATTAAGGATATCTGTATATCTTTGGAAAGTAC	563
OY	676	TTTTCGATTATGACCACCAAAAAAATTTGGTGGAGAT---AAAAAATGGGGGCTTGGCTCAC	732
Db	564	TTTTTCAGATCTTACCGCATCAAAAATTTGGTAGTAATTTGCGCGGCACCTGAGCTTGTGCTC	623
OY	733	TAACTGAGACACTTTTGGCAAGAGCTCTGGAAGCCAAACAGAAATCAGGCTTTGGGAAT	792
Db	624	TGCAGAGGAGGTTATTTTGGCAAGAGCTTCGAAGCCCAAGAAATTAAGGCTGGAAACT	683
OY	793	GGGAGTTTATTCCTGTGCA	813
Db	684	GGGAGTTATATCTTTATCA	704

RESULT 4	
AM201789	
LOCUS	
DEFINITION	AM201789 552 bp mRNA linear EST 02-DEC-2001 sf07c10.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl027-1579 5' similar to TR:Q92242 Q92242 WRN PROTEIN.; , mRNA sequence.
ACCESSION	AM201789
VERSION	AM201789.1 GI:6482532
KEYWORDS	EST.
SOURCE	soybean.
ORGANISM	Glycine max
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine. 1 (bases 1 to 552)
AUTHORS	Shoemaker, R., Kaim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theisling, B., Allen, M., Bowers Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk R., Riltter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann R., Waterston, R. and Wilson, R. Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project
TITLE	
JOURNAL	
COMMENT	

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ReesGen, Invitrogen Corp. 2130

Db 515 TTTCAGATTAGCCACCAAAAAATTTGGTGAAGATATAAAAAATGGGCGCTTCCTCACTAA 574

Qy 736 CTGAGACACTGTTGTTGCAAGAGCTCTGAAGCCAAACAGATCAGCTTGGGAAGTGGG 795
|||||
Db 575 CTGAGACACTGTTGTTGCAAGAGCTCTGAAGCCAAACAGATCAGCTTGGGAAGTGGG 634

Qy 796 AGTTTATCCTCTGTCAAGCAGATTACATACGC 832
|||
Db 635 AGGTCATNCTCTATCAAGCAGCAGTACATNCGC 671

RESULT 6
AM832139 467 bp mRNA linear EST 03-DEC-2001
LOCUS
DEFINITION sm20b11.y1 Gm-c1027 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1027-9430 5' similar to SW:WRN_HUMAN Q14191 WERNER SYNDROME
HELICASE. [1] ; , mRNA sequence.
ACCESSION AM832139
VERSION AM832139
KEYWORDS EST.
SOURCE soybean.
ORGANISM glycine max
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eustosids I: Fabales: Fabaceae: Papilionoideae: Phaseoleae:
Glycine.
1 (bases 1 to 467)
Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Coryell, V., Khanna
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuk
, R., Riter, E., Kohn, S., Shu, T., Jackson, Y., Cardenas, M., Mccann
, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccuteresgen.com
Insert Length: 734 Std Error: 0.00
High quality sequence stop: 417.
Location/Qualifiers
1.467
/organism="glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1027-9430"
/clone_lib="Gm-c1027"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-stranded synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGAGACTACTCTCGAG(T)18] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and

subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript(tm) II XR predigested vector
(pBluescript II SK(+)) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). 97% of the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=30). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Coryell."

BASE COUNT 133 a 84 c 109 g 140 t 1 others
ORIGIN

Query Match 23.0%; Score 241.2; DB 10; Length 467;
Best Local Similarity 70.0%; Pred. No. 1,6e-49;
Matches 324; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy 430 TTGGCTTGATATTGAGTGGAGACCAAGTTTAGAAAAGGTCTCCCGGGAAGTTG 489
|||||
Db 4 TTGGATTGGACATTGATGGAACCCACCTTCAGAAAAGGTCTCCCGGAAGTAG 63

Qy 490 CGACTGCCAGATATGTGTAGATAGTAATTATTTGATGATTATCTTTCATTCTG 549
|
Db 64 CAGTGATGCAATATGTGTGACACTAGACATTTGCTTCATCTATTCATTCATTCTG 123

Qy 550 GTATCCCTCAAGACGTCCCAACATCTTATGAAAGTTCAACACTGTAAAGTAGTATG 609
|
Db 124 GAATCCCTCAAAATTTACAGCTTTTGGTTGTAAGATCCCAAGCTTTAAAGGTGAGCT 183

Qy 610 GAATTGATGTGACTGTGCAAGCTTTTCATGACTGTAGTATGATCAAGATGTTG 669
|
Db 184 GGATTATGGATGATGCTGTAGAGTTTATGAGATTTATCAATATCTGTTAAAGTGTGA 243

Qy 670 AGGATCTTTACAGATTAGCCAAACCAAAATTTGGTGAAGATAAAAATGGGCTTCCT 729
|
Db 244 CGGATCTTTCTTTTCATGCTATCAAAAGCTTGGTGAAGATCAATAGTGGGCTTCCTCAT 303

Qy 730 CACTTACTGAGACACTGTTTGGCAAGAGCTCCGAGCCCAACAGATACAGCTTGGGA 789
|
Db 304 CTTGACTGAAAACCTCTATCAACAGCTTAAAGGCCCAACAAATTAAGACTGGAA 363

Qy 790 ACTGGAGTTTATCTCTGTCAAGCAGAGTATCAATACGCAAGCAGATCTTATG 849
|
Db 364 ATTGGAGGCTCCGTTTGTCAAGGAGCAGACATAGATTCGCAACAGATCTTTTG 423

Qy 850 CTTGATGGCATCTTTACAGAGTTCTTAAGGACCTTCTGATGC 892
|||||
Db 424 CTTCTTGTTGTCTTATCAGCGCATTAAGATCTCCGAGAGC 466

RESULT 7
BQ799068
LOCUS
DEFINITION BQ799068 661 bp mRNA linear EST 30-JUL-2002
EST 1237 Green Grape berries lambda zap II library Vitis vinifera
BQ799068
ACCESSION BQ799068
VERSION BQ799068.1 GI:22014034
KEYWORDS EST.
SOURCE Vitis vinifera.
ORGANISM Vitis vinifera.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Vitaceae: Vitis.
1 (bases 1 to 661)
Abbal, P., Agasse, A., Ageorges, A., Atanassova, R., Barrieu, F.,
Couture, C., Dedalechamp, F., Delrot, S., Glissant, D., Grimplet, J.,
Hamdi, S., Romieu, C. and Terrier, N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages

RESULT 9
LOCUS BG133220 506 bp mRNA linear EST 31-JAN-2001
DEFINITION EST466112 tomato crown gall Lycopersicon esculentum cDNA clone
coto11b12 5' sequence, mRNA sequence.
ACCESSION BG133220
VERSION BG133220.1 GI:12633408
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eunasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 506)
van der Hoeven,R., Sun,H., Cho,J., Uutterback,T., Hansen,C., Romning
C. and Tanksley,S.
Generation of ESTs from tomato crown gall tissue
Unpublished (2001)
JOURNAL Contact: CGI
COMMENT Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
FEATURES
source
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="coto11b12"
/clone_lib="tomato crown gall"
/tissue_type="crown gall"
/dev_stage="crown galls from full-grown plants (8 wks old)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Four wk old greenhouse plants were stab inoculated
on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,
Cornell U.). Galls were allowed to develop for another 4
wks, when gall tissue was frozen in liquid nitrogen."
BASE COUNT 142 a 91 c 126 g 147 t
ORIGIN
Query Match 21.3%; Score 223; DB 12; Length 506;
Best Local Similarity 65.2%; Pred. No. 5; 6e-45;
Matches 328; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
QY 363 GATACCGAGCAATGACGCTTATTAAGTTCTTGATACCAAGAGATGATCTGCATA 422
DB 2 GAGAAATCTGCAAGTAGAGCTGTTAAATTTGTTGAGGAAAGAGAAAGAGGCAAT 61
QY 423 GCTTTGGTGGCTGGATATGATGAGAGACCAAGTTTGAAGAGGTGTTCCCGGGG 492
DB 62 GTTGCTCTTGATTTGACATTTGATGAGAGCCACCTTTTGAAGAGGTGCGACCTGG 121
QY 483 AAGGTGCGACTGTCAGATATGTAGATAGTAAATTTATGATGATTTATGATATTTT 542
DB 122 AAGGCTGCTGTTATGAGATATGTGTGACAAAGGTAATTTGATTTGATATCATC 181
QY 543 CATTTGATGATCCCTCAAGTCTCCAAACATCTTATTTGAAGATTTCAACACTTTGA 602
DB 182 CACTCGGAGATCCCTCAACTCTGCAATCTCTTCTGAGATTCACACTGTTGGAAGTG 241
QY 603 GGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
DB 242 GGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
QY 663 GATGTTGAGATCTTACAGATTAGCCCAACCAAAATTTGTTGAGATGATGATGATGAT 722
DB 302 GCTTTGGAAGATCTTCCGACACTTGCACAAACAAACCTTGATGATGATGATGATGAT 361

QY 723 CTTCCTCACTAAGTACGACACTGTTTTCGAAGAGCTCTGAGACCAACAGATCAG 782
DB 362 CTAGCATCTTAAGTACGAGAGCTTCTTGCCAGACAGCTCCCAAGGCAATATATCAG 421
QY 783 CTGGGAGACGAGAGTTTATCTCTGTCGAACAGACAGTATCAATACGACAGAGAT 842
DB 422 TTGGGGAATTTGGAGGCTATGTTATCTAGGACCACTACATTAATGCTGATCAGAT 481
QY 843 GCTTATGCTTCATGCGATCTTTA 865
DB 482 GCCTTGTGTTCCGTGACTATATA 504
RESULT 10
LOCUS BM094501 531 bp mRNA linear EST 30-NOV-2001
DEFINITION sa116903.y1 Gm-cl066 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl066-3005 5' similar to TR:Q9SYM6 Q9SYM6 HYPOTHETICAL 35.1
KD PROTEIN.; mRNA sequence.
ACCESSION BM094501 GI:17023467
VERSION BM094501.1
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 531)
Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Correll,V., Khana
A., Bolla,B., Matra,M., Hillier,L., Kuehba,T., Matlin,J., Beck,C.,
Wyle,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Rilter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST project
Unpublished (1999)
JOURNAL Contact: Shoemaker R./Public Soybean EST Project
COMMENT Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntleville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cout@resgen.com
High quality sequence stop: 423.
FEATURES
source
Location/Qualifiers
1..531
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-3005"
/clone_lib="Gm-cl066"
/tissue_type="leaf and shoot tip, salt stressed, 2 week
old seedling"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from mRNA isolated
from unexpanded leaves and the shoot tips of 2 week old
seedling from the cultivar Williams. The 2 week old
seedlings were salt stressed in a solution of 500mM NaCl
for 3 days prior to harvesting. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

ORIGIN

Query Match 20.4%; Score 213.6; DB 13; Length 531;
Best Local Similarity 64.6%; Pred. No. 1.2e-42;
Matches 340; Conservative 0; Mismatches 174; Indels 12; Gaps 1;

OY 299 TAGAATTTTCCAGCATAGTTGGTGTAGAGATTTTGTATAGCAGACAGTCTGCTACTGA 358
16 TATGACATGTCCTCTATAGTGAAGTTAGTGTGCTCAAAATTTTATATATCAGACCTTTGATGC 77
DB 355 GGTGATTAAGCAGCATGACGCTTATTAAGTTCTTGATATACAGAGAGATGAATCTGG 418
78 TGTAGAGAAAGCTGCAACAAGCTTACAAATTTCTCAAGATGATGCA----- 128

OY 419 AATAGCTTTTGTGGCTTGATATTTGAGTGAAGACAGTTTATAGAAAAGCTTCTCC 478
129 ---AATTGCAATTTGATTTGATGATGTAAGAACCCACCTTGAGAAAAGCTTCTCCACC 185

OY 479 GGGGAGATTTGCGACTCTCCAGATATGTGATAGTAAATTTATTTGATGATGATAT 538
186 CGGAAAGGTAGCAGATGATGATATATGATGACAGTATGTCATCTTCTACATCT 245

OY 539 TTTTCAATCTGATATCCCTCAAGTCTCCACATCTTATTTGAAGATTTCAACACTGTAA 598
246 AATTCAATCTGGAATCCCTGGAATTTACAGCTTTTGTGGAAGATCCACAGCTCTGAA 305

OY 599 GGTAGGATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
306 GGTGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365

OY 659 CAAAGATTTGAGATCTTTCAGATTTAGCCAAACAAAATTTGATGATGATGATGAT 718
366 TAAAGGTGACGAGATCTTCTTTTCATGCTATTCGAAGCTTGTGAGATCATATAGTG 425

OY 719 GGGCTTTGCTCTACTAGTGAAGACTGTGTTGCAAGAGCTCTCTAAGCCAAACAGAT 778
426 GGTGCTTCTGATCTTGTAGCAAAACCTTTATATCAAAACCTTAAAGCCCTAACAAT 485

OY 779 CAGGCTTTGGAAGCTGAGATTTATCTCTGTCGAAGCAGCAGCTTA 824
486 AAGACTGGGAATTTGGAGACTCTCTGTTTGTCAAAAGAGCAACTA 531

DB 486 AAGACTGGGAATTTGGAGACTCTCTGTTTGTCAAAAGAGCAACTA 531

RESULT 11
BM358892 515 bp mRNA linear EST 09-JAN-2002
LOCUS GA_Ea0013P02r Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION
ACCESSION BM358892
VERSION BM358892.1 GI:18099638
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 515)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
JOURNAL Contact: Wing RA
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total High Quality bases = 415
Seq primer: TAATAGCACTCACTAATAGG
High quality sequence stop: 514.
Location/Qualifiers

source 1.. 515
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0013P02r"
/clone_1b="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBR-CMW; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 134 a 112 c 109 g 160 t

Query Match 16.4%; Score 172.2; DB 13; Length 515;
Best Local Similarity 64.6%; Pred. No. 2.6e-32;
Matches 277; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

OY 208 ATCAATCCCAATTAATCCGTCGCAATTCGCTGTCATCATCTTCTACATCTT 267
87 ATCAAGATTAACCTCAAAACAGCGCCATTCGCCATTCATCTCTACCGCCGCTT 146

OY 268 ATAAAGATTTCTCTCTCCGTTGCCAGCTA-----GGAATTTCCAGCATGA 318
147 CTTTCCGTCCTTCTCCGCTCGCTGCAATCTAATCAAGGCTTATGATATCCGCCCTTGA 206

OY 319 GGTGCTGCTAGATTTTGTATAGCAAGCTCTGAGCTGATATAGCGAGCAATGC 378
207 GATTTGAGGCTATATTTGTATAGCTTTAGCGAAGATGAGGAGAAATGCTGCATGG 266

OY 379 AGCTTTTAATAGTCTTGTATACCAAGAGATGAATCTGGAATAGCTTTTGTGCTTG 438
267 AGCTATTAATAATTTGTGAATTAAGAAAGAAATGAGTCAAGTTGCTTAAGGTTTG 326

OY 439 ATATGAGTGGAGACCAAGTTTGAAGAGTCTTCCCGGGAAGTTGGAGCTGTC 498
327 ATATGAGTGGAGAGCCCTCTTCCAAAGGAATTTTGGCTGGAGGCTGGGCTGATGC 386

OY 499 AGATATGTGATAGTAAATATTTATGATGATGATGATGATGATGATGATGATGATG 558
387 AGATATGTGATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 446

OY 559 AAAGTCTCAACATCTTATTTGAAGATTTCAACACTTTAAAGGTAGTATTTGAATTTGAT 618
447 AAAGTCTGACAGTCTTCTTGAAGACTCCGAATTTATAAAGTTGGAGTTGCAATTTGATG 506

OY 619 GTGACTCTG 627
507 GCGATGCTG 515

DB 507 GCGATGCTG 515

RESULT 12
BM358891 517 bp mRNA linear EST 09-JAN-2002
LOCUS GA_Ea0013022r Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION
ACCESSION BM358891
VERSION BM358891.1 GI:18099637
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 517)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
JOURNAL Contact: Wing RA
COMMENT Clemson University Genomics Institute

Oy	531	ATGCATTATTTTCCTGATCGTACCCCAAGCTCCAAACATCATTAATGAGATTCACA	590
Dd	406	CTACACTTAATTCATTTCTGGAAATCCCTCAAATTTACAGCTTTTGCTTGGAATCCCA	465
Oy	591	CTTGTAAGGTAGATTAATGGAATGATGATGCTGTCGTAAGCTTTTCCATGACTATGGA	650
Dd	466	GTCCTTAGAGGTTGAGCTGGGATTCATGCTGTGATGCTGGAAGCTTTTAGAGATTATAAC	525
Oy	651	GTTAGTATCAAGAATGTTGAGGATCTTTC	679
Dd	526	ATATCTGTTAAAGGTGACGAGATCTTTC	554
RESULT 14	Bj213300	595 bp	mRNA linear EST 04-APR-2002
LOCUS	Bj213300		
DEFINITION	Bj213300 Y. Ogihara unpublished cDNA library, wh Trilicium aestivum		
ACCESSION	CDDA Clone wh2lm08 5', mRNA sequence.		
VERSION	Bj213300		
KEYWORDS	Bj213300.1 GI:19952807		
SOURCE	EST.		
ORGANISM	bread wheat.		
JOURNAL	Trilicium aestivum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae		
TITLE	1 (bases 1 to 595)		
JOURNAL	Triliceae; Trilicum.		
COMMENT	Ogihara,Y and Murai,K Expressed genes in Trilicum aestivum Unpublished (2002) Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp. location/Oallifiers		
FEATURES	source		
	1..595		
	/organism="Trilicum aestivum"		
	/cultivar="Chinese Spring"		
	/db_xref="taxon:4565"		
	/clone="wh2lm08"		
	/clone_id="Y. Ogihara unpublished cDNA library, Wh"		
	/ltsuse_type="spike at meiosis"		
	/dev_stage="Peekes' scale 9"		
	/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown under hydropnic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhnov in J Dvorak lab). Total RNA was prepared from sheath tissue, equal quantities of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pluescript phagmids in the T3 Close lab at the University of California, Riverside (Akhnov, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."		
BASE COUNT	170 a 119 c 148 g 157 t		1 others
ORIGIN			
Query Match	15.4%:	Score 162:	DB 13: Length 595:
Best Local Similarity	58.3%:	Pred No 9,le-30:	
Matches 323: Conservative	0: Mismatches 225:	Indels 6:	Gaps 2:
Oy	321	TTTTGGTGTAGCATTTCATAGCAAGACCTCTACTGAGTTGATTAACGACGATCGAG	380
Dd	45	TTCACTGCGCAGATAGTCTACTGCCGCGGCGCATTAAGTGAAGAAAGCTACGCGGAG	104
Oy	381	CTTATTAAAGTTCTTGATACCAAGAGATAATCTGCAATAGCTTTTGTGGCTTGAT	440

	DB	105	ATTGACGGAAATGCAGAACATGA - --AGGCCCTCGGGCCGGTCCCTTCGTTCAT	161
	OY	441	ATTGAGTGAGACCACCAAGTTTAGAAAAGSTGTTCTCCGGGAGAAGTTGCCACTCAG	500
	DB	162	CTCGAGTGAAAAACCTTTCCCGAAGAGAGAACCCCATGTAAGTCGCTGATGCNA	221
	OY	501	AATATGTAAGATAATTAATTATGATGTAATGACAATTTTTTCATTCGTATCCCTCA	560
	DB	222	TATATCATGACGACAAACATCATATTGATATCTCATGCAATATCATCTCTGGCGTCCC	281
	OY	561	AGTCGCCAACATCTTATTGGAAGATTCAAACACTTTGTAAGGTAGTATTTGAAATGATGT	620
	DB	282	ATCTTGAAATATCTCTTTTGAGAGACAGTTTCATCCGTTAAAGTTGGAGTATGATACAAAT	341
	OY	621	GACTCTGTGAACCTTTTCCATGACTATGAGAGTTAGTATCAAAAGATGTTGAGATCTTTCA	680
	DB	342	GATGCAAGGAAANAATGTTCAATGATATGATGATGCCGTGTACAACCATATGATGATTTARCA	401
	OY	681	GATTATGCAACCAAAAAAAAAATTTGCTGSG ---AGATAAAAAATGGCGCTTGCCCTACTACT	737
	DB	402	ACTGTTGCAAAACGTCACAGTTAGCTGGGCCCTATTAATAAGATGAGCTTGTGTCATTAATCT	461
	OY	738	GAGACACTTGTTTGGCAAGAGCTCCTGGAAGCAAAACAGATCAGGCTTGGAGATGGAG	797
	DB	462	GAAATGCTTACTATGAAGAGAGTTGGCCAACGCTGGCAACATTAAGATGGAACACTGGAG	521
	OY	798	TTTATCTCTGTGTCAAGACGACGATTACAAATATGSCAGCACAGCATGCTTATGCTTCATGG	857
	DB	522	TCTTTTGTTCTCTCAAAAAGCACTGATGATGTGCTGTACGAGATCCTACATCTCTGG	581
	OY	858	CATCTTTTCAAGGT 871	
	DB	582	TACTTTGATGAGGT 595	
RESULT_15	BIA33038			
LOCUS	BIA33038			
DEFINITION	BIA33038	499 bp mRNA linear EST 21-AUG-2001		
ACCESSION	E57535799 P.	Infestans-challenged leaf Solanum tuberosum cDNA clone		
VERSION	PICAZ86.5	sequence, mRNA sequence.		
KEYWORDS	BIA33038.1 GI:15257728			
SOURCE	EST.			
ORGANISM	potato.			
	Solanum tuberosum			
	Eukaryota; Viridiplantae;	Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;			
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum.			
REFERENCE	1 (bases 1 to 499)			
AUTHORS	Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chelmingo,A., Bongti,O., Buell,C.R., Roming,C.M., Fry,W.E. and Baker,B.			
TITLE	Generation of ESTs from Potato Leaves Challenged with Phytophthora infestans, Compatible Interaction			
JOURNAL COMMENT	Unpublished (2000) Contact: Cathy Romning The Institute for Genomic Research For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com Seq primer: M13F-R.			
FEATURES				
source	Location/Qualifiers			
	1..499			
	/organism="Solanum tuberosum"			
	/cultivar="Kennebec"			
	/db_xref="taxon:4113"			
	/clone="PICAZ86"			
	/clone_lib="P. infestans-challenged leaf"			
	/tissue_type="leaf"			
	/dev_stage="6 week old"			
	/lab_host="SOLR"			
	/note="Vector: plasmid SK(-); Site 1: EcoRI; Site 2: XhoI; Whole plants were challenged with 450,000 sporangia/ml P. infestans US-1(IUS 940501) in Biotron (Madison, Wisconsin). Leaf tissue was collected at 1, 2,			

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 15:06:14 ; Search time 52 Seconds

(without alignments)
6720,758 Million cell updates/sec

Title: US-09-896-186b-23

Perfect score: 1049
Sequence: 1 accaagcattatatttattt.....tttaacgctcagaactag 1049

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 310279 seqs, 166577418 residues

Total number of hits satisfying chosen parameters: 620558

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PC1_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PC1US_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	10.6	5189	US-09-954-456-1131	Sequence 1131, Ap
2	44.4	4.2	3149	US-09-729-674-127	Sequence 127, App
3	38.2	3.6	439	US-09-864-761-20174	Sequence 20174, A
4	36	3.4	99916	US-09-816-095-3	Sequence 3, Appli
5	36	3.4	659158	US-09-771-208-20	Sequence 20, Appli
6	35.8	3.4	29544	US-09-464-767-1	Sequence 1, Appli
7	35.8	3.4	32745	US-09-464-767-3	Sequence 3, Appli
8	35	3.3	954	US-09-947-971-3	Sequence 3, Appli
9	34.6	3.3	280	US-09-933-797-413	Sequence 413, App
10	34.4	3.3	463	US-09-864-761-5350	Sequence 5350, Ap
11	34.4	3.3	53326	US-09-818-264-3	Sequence 3, Appli
12	34	3.2	527	US-09-864-761-6599	Sequence 6599, Ap
13	34	3.2	527	US-09-864-761-8066	Sequence 8066, Ap
14	34	3.2	1827	US-09-925-302-270	Sequence 270, App
15	33.8	3.2	556	US-09-864-761-8619	Sequence 8619, Ap
16	33.8	3.2	785	US-09-770-445-860	Sequence 860, App
17	33.6	3.2	258	US-09-878-574-10134	Sequence 10134, A
18	33.6	3.2	285	US-09-878-574-7317	Sequence 7317, Ap
19	33.6	3.2	5814	US-09-764-847-1860	Sequence 1860, Ap

ALIGNMENTS

C	20	33.4	3.2	533	10	US-09-864-761-6759	Sequence 6759, Ap
	21	33.4	3.2	2979	10	US-09-815-242-4754	Sequence 4754, Ap
	22	33.4	3.2	3009	10	US-09-815-242-8801	Sequence 8801, Ap
	23	33.4	3.2	3009	10	US-09-815-242-9060	Sequence 9060, Ap
C	24	33.4	3.2	7035	10	US-09-815-242-8615	Sequence 8615, Ap
	25	33.2	3.2	477	10	US-09-864-761-5436	Sequence 5436, Ap
	26	33	3.1	305	10	US-09-864-761-18262	Sequence 18262, A
	27	33	3.1	496	10	US-09-864-761-2534	Sequence 2534, Ap
C	28	33	3.1	1671	12	US-10-052-586-505	Sequence 505, App
	29	32.8	3.1	3294	10	US-09-764-860-1156	Sequence 1156, App
C	30	32.8	3.1	203654	10	US-09-820-905-3	Sequence 3, Appli
	31	32.4	3.1	471	10	US-09-795-668-1424	Sequence 1424, Ap
	32	32.4	3.1	471	10	US-09-795-668-1424	Sequence 1424, Ap
	33	32.4	3.1	1038	10	US-09-815-242-9161	Sequence 9161, Ap
	34	32.4	3.1	66109	10	US-09-880-107-3768	Sequence 3768, Ap
	35	32.4	3.1	198285	10	US-09-880-107-3814	Sequence 3814, Ap
C	36	32.4	3.1	640681	10	US-09-790-988-1	Sequence 1, Appli
	37	32.2	3.1	1956	10	US-09-351-794A-1	Sequence 1, Appli
C	38	32.2	3.1	5340	10	US-09-817-360-1	Sequence 1, Appli
	39	32.2	3.1	21860	10	US-09-764-877-3803	Sequence 3803, Ap
C	40	32	3.1	153	10	US-09-924-035A-86	Sequence 86, Appli
	41	32	3.1	575	10	US-09-864-761-8070	Sequence 8070, Ap
C	42	32	3.1	1185	10	US-09-738-363-9	Sequence 9, Appli
	43	32	3.1	1185	10	US-09-977-653-9	Sequence 9, Appli
	44	32	3.1	1237	10	US-09-070-927A-555	Sequence 555, App
C	45	32	3.1	1577	10	US-09-864-761-30755	Sequence 30755, A

RESULT 1
US-09-954-456-1131
Sequence 1131, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1131
LENGTH: 5189
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1131
Query Match 10.6%; Score 111; DB 10; Length 5189;
Best Local Similarity 54.1%; Pred. No. 2.4e-22;
Matches 272; Conservative 0; Mismatches 225; Indels 2;

OY	336	GATACCAAGACAGATGATCTGGAATAGCTTTTGTGGCTTGATTTGGATGGACCA	4555
Db	433	GATATTACCATGAGCTCTATCAAGATGGGATGTGGTGGATTTGCATGAGATGGCACC	4222
OY	456	AGTTTTCAGAAAAGTGTCTCCGGGGGAAGGTGCGACTGCGAGATATGTAGATAGT	5151
Db	493	TTATTCAAATTAAGGGAACCT---TGCGAAAGTTGCACATTAATTCAGTTGTGTGTTGAG	5499
OY	516	AATTATGTGATGTATGACATATTTTT---CATCTGCTATCCCTCAAGCTTCCACAT	5727
Db	550	AGCAAAATTTTACTTTGTTCCAGCTTTCTTCCATGTCAATTTTTTCCCGAGGATTTAAATATG	6090
OY	573	CTTATTCAGATTCACACCTGTAAAGTAGTATTGGAATGTGATGCTGTGTAAG	6322
Db	610	TTGCTTGAATAATTAACGACGTTAAATAAGCAGCTGATGAATTTGAAGAGATCAAGTGAA	6659
OY	633	CTTTTCCATGACTATGGAGTTAGTATCAAGATTTGAGATCTTTTACATTTAGCCAC	6522
Db	670	CTTCTACGTGACTTTGATATCAAAATTTGAAGATTTTGTGAGTTGACAGATGTTGCCAAT	7299
OY	693	CAAAAATTTGGTGGAGATAAAAAATGGGCGCTTGCCCTCACAACTGAGACACTTGTTCG	7522
Db	730	AAAAAGCTGAATGTACAGNAGACCTGGAGCCTTAAACAGTCTGTTTAACACCTTTAGGT	7899
OY	753	AAAGAGCTCTGAAAGCCAAACAGATCAGGCTTGGGAACCTGGAGTTTATCCTCTGTCA	8122
Db	790	AAACAGCTCTCGAAGAACCAACACTATCCGCTGTGACATTTGGAGATAAATTTCCCTCCACT	8499
OY	813	AAGCAGAGTTACATATACGACAGCAAGGATGCTTATCTTCATGAGCATCTTTACAAAGTT	8722
Db	850	GAGGACCGAAGAACTGTATGACGACCACTGATGCTTATCTGTTTTATTATTACCGAAAT	9099
OY	873	CTTAAGGACCTTCCGTGATGCTGT	895
Db	910	TTAGAGATTTTGGATGATACGT	932

RESULT 2
 US-09-729-674-127
 Sequence 127, Application US/09729674
 Patent No. US2001003935A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 APPLICANT: McCoy, John M.
 APPLICANT: Lavallie, Edward R.
 APPLICANT: Collins-Racie, Lisa A.
 APPLICANT: Evans, Cheryl
 APPLICANT: Merberg, David
 APPLICANT: Treacy, Maurice
 APPLICANT: Agostino, Michael J.
 APPLICANT: Steinginger II, Robert J.
 APPLICANT: Spaulding, Vlki1
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fechtel, Kilm
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 FILE REFERENCE: 6055-64X
 CURRENT APPLICATION NUMBER: US/09/729,674
 CURRENT FILING DATE: 2000-12-04
 PRIOR APPLICATION NUMBER: 09/559,330
 PRIOR FILING DATE: 2000-03-30
 NUMBER OF SEQ ID NOS: 283
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 127
 LENGTH: 3149
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-729-674-127

Query Match 4.28; Score 44.4; DB 10; Length 3149;
 Best Local Similarity 49.3%; Pred. No. 0.0069;

[illegible]

```

1  RESULT 3
2  US-09-864-761-20174/c
3  ; Sequence 20174, Application US/09864761
4  ; Patient No. US20020048763A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Penn, Sharron G.
7  ; APPLICANT: Rank, David R.
8  ; APPLICANT: Hanzel, David K.
9  ; APPLICANT: Chen, Wenheng
10 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
11 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
12 ; FILE REFERENCE: Aeomica-X-1
13 ; CURRENT APPLICATION NUMBER: US/09/864,761
14 ; CURRENT FILING DATE: 2001-05-23
15 ; PRIOR APPLICATION NUMBER: US 60/180,312
16 ; PRIOR FILING DATE: 2000-02-04
17 ; PRIOR APPLICATION NUMBER: US 60/207,456
18 ; PRIOR FILING DATE: 2000-05-26
19 ; PRIOR APPLICATION NUMBER: US 09/632,366
20 ; PRIOR FILING DATE: 2000-08-03
21 ; PRIOR APPLICATION NUMBER: GB 24263.6
22 ; PRIOR FILING DATE: 2000-10-04
23 ; PRIOR APPLICATION NUMBER: US 60/236,359
24 ; PRIOR FILING DATE: 2000-09-27
25 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
26 ; PRIOR FILING DATE: 2001-01-30
27 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
28 ; PRIOR FILING DATE: 2001-01-30
29 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
30 ; PRIOR FILING DATE: 2001-01-30
31 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
32 ; PRIOR FILING DATE: 2001-01-30
33 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
34 ; PRIOR FILING DATE: 2001-01-30
35 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
36 ; PRIOR FILING DATE: 2001-01-30
37 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
38 ; PRIOR FILING DATE: 2001-01-30
39 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
40 ; PRIOR FILING DATE: 2001-01-30
41 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
42 ; PRIOR FILING DATE: 2001-01-30
43 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
44 ; PRIOR FILING DATE: 2001-01-30
45 ; PRIOR APPLICATION NUMBER: US 60/234,687
46 ; PRIOR FILING DATE: 2000-09-21
47 ; PRIOR APPLICATION NUMBER: US 09/608,408
48 ; PRIOR FILING DATE: 2000-06-30
49 ; PRIOR APPLICATION NUMBER: US 09/774,203
50 ; PRIOR FILING DATE: 2001-01-29
51 ;

```

```
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
SEQ ID NO 20174
LENGTH: 439
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035419.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
US-09-864-761-20174

Query Match
Best Local Similarity 43.9%; Score 38.2; DB 10; Length 439;
Pred. No. 0.14;
Matches 163; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

QY 340 ATAGCAAGACTGCTACTGAGTTGATTAACCGAAGCAATGACCTTATTAAGTTCTTGATA 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 374 ATGGGATGCTCTGCTGATGATGATGGGATTAATGATGATGCTGCTCGATGGTG 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 CCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 314 ATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 TTAGAAAAGTGTTCCTCCGGGGAAGTTCGACATGCTCCAGATGATGATGATGATGAT 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 254 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 GTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 AAGATTCAACACTTGAAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 GTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 74 AATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 TTGGTGGAGAT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14 ATGGTGATGAT 4
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-09-816-095-3/C
Sequence 3, Application US/09816095
Patent No. US20020137164A1
GENERAL INFORMATION:
APPLICANT: GAN, WEI-LIU
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
TITLE OF INVENTION: THEROPF
FILE REFERENCE: C1001147
CURRENT APPLICATION NUMBER: US/09/816,095
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 99916
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
```

```
LOCATION: (1)...(99916)
OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match
Best Local Similarity 46.2%; Score 36; DB 10; Length 99916;
Pred. No. 17;
Matches 120; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 425 TTTTGTGGCTGGATATTGAGTGGACCAAGTTTGAAGAAAGTGTCTCCGGGGA 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41649 TCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 485 GGTGGGACTGTCACATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41589 GAATTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 545 TTCTGTATCCCTCAAGTCTCCACATCTTATTAAGATTCACCTGTAAAGTAG 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41529 TGCCCTAATTTCAAGAGCTGTCTGATGATGATGATGATGATGATGATGATGATGAT 41470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 605 TATTGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41469 TCTTGGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 41410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 665 TGTGAGATCTTTCAGATT 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41409 TCGGTAGAGGTGTTTATGAT 41390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-09-771-208-20
Sequence 20, Application US/09771208
Patent No. US20020155564A1
GENERAL INFORMATION:
APPLICANT: MEDRANO, JUAN
APPLICANT: BRADFORD, ERIC
APPLICANT: HORVAT, SIMON
TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
FILE REFERENCE: 407P-923710US
CURRENT APPLICATION NUMBER: US/09/771,208
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: US 08/999,477
PRIOR FILING DATE: 1997-12-29
SOFTWARE: PatentIn version 3.0
NUMBER OF SEQ ID NOS: 20
SEQ ID NO 20
LENGTH: 659158
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (123459)...(123478)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (602466)...(602485)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (546998)...(547017)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (494715)...(494814)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (390986)...(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (346860)...(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (317174)...(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (280353)...(280373)
```


; Sequence 8066, Application
; Patent No. US20020048763A1


```

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8066
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ACO13707.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; US-09-864-761-8066

Query Match          3.2%; Score 34; DB 10; Length 527;
Best Local Similarity 50.6%; Pred. No. 2.7;
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Oy 531 ATGCAATATTTTCATTCGTATCCCTCAAGATCTCCAAACATCTTATGGAATTCACA 590
    || || || || || || || || || || || || || || || || || || || || ||
Db 226 ATCCACATATTTTAAATGGTGAAGATATATAGCTATATATATATATATTCACCAAGTAA 167

Oy 591 CTTTGAAGGTAGTATTTGGAATTTGATGTGACTCTGTGAAGCTTTTCCATGACTATGGA 650
```

```

; RESULT 14
; US-09-925-302-270/c
; Sequence 270; Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 270
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-302-270

Query Match          3.2%; Score 34; DB 10; Length 1827;
Best Local Similarity 56.1%; Pred. No. 5.7;
Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Oy 10 TTAATTTTATTTTGTTCAGTAAAGAAATGTCATCGCAATGATGAGACGAG 69
    || || || || || || || || || || || || || || || || || || || || ||
Db 1808 TTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1749

Oy 70 CTTTACAGAGGAAGACTTCCTGCTATGACGCGCATCGAAGCTTCTCAATTT 123
    || || || || || || || || || || || || || || || || || || || || ||
Db 1748 ACTGTGCCCGCAAGATGCTCTCCAAACATGSAACCAAGAAACAGTAAATTT 1695

; RESULT 15
; US-09-864-761-8619
; Sequence 8619; Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```

PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 8619
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC023000.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-8619

```

Query Match	3.2%;	Score 33.8;	DB 10;	Length 556;
Best Local Similarity	46.7%;	Pred. No. 3.2;		
Matches 107; Conservative	0;	Mismatches 122;	Indels 0;	Gaps 0;

[illegible]

Search completed: November 2, 2002, 16:42:44
Job time : 788 secs

Db 433 GATATTAGCATGATCTATCAGATGGGATGTGGGATTTGCATGGATGGCCACCA 492
QY 456 AGTTTGAAGAAAGTGTCTCCGGGAGAGTTGCGACTGCGCATGATGTGATAGT 515
Db 493 TTATACATATGAGGGAAGT---TGCGAAGTTGCTACTAATTCAGTGTGTTCTGAG 549
QY 516 AATTATTGTGATGTATGATATTTT---CATTTGTGTATCCCTCAAGTCTCCAAAT 572
Db 550 AGCAAAATGTACTGTTCACAGTTTCTTCATGTCAGTTTTCCTCCAGGATTAAGT 609
QY 573 CTATTTGAAGATTCACATCTTTGAAGTGTGTAAGTGTGATGTGATGCTGTGAAG 632
Db 610 TTGCTTGAAGATTAAGGATTAAGGAGGAGGTGTAGAAATGGAATGAGATCAGTGAAA 669
QY 633 CTTTTCATGACTATGAGTATGATGATCAAGATGTGAGATCTTTGATGATTTAGCCAA 692
Db 670 CTTTACGTGACTTGTGATATCAAAATTTTGTGAGTTGACAGATGTTGCCAAT 729
QY 693 CAAAAAATGTGTGAGATTAAGAAATGGGGCTTGCTCCTACATACCTGAGACATGTTTGC 752
Db 730 AAAAAGCTGAATGTACAGAGACCTGGACCTTAAACAGTCTGGTTAAACACCTTTAGT 789
QY 753 AAAAGCTCCTGAAGCCAAACAGATCAGGCTTGGAACTGGAGTTTATCCTCTGTCA 812
Db 790 AAACAGCTCCTGAAAGACAGATCTATCCGCTGTAGCAATTAATTTCTCTCACT 849
QY 813 AAGCAGATTTACATACGAGCAACGATGCTATGCTTCATGCTTACAGATTTTACAGGTT 872
Db 850 GAGACCCAGAAACGTGTGACAGCCACTGATGCTTATGCTGTTTATTTATTTACCGAAT 909
QY 873 CTTAAGGACCTTCCCTGATGCTGT 895
Db 910 TTAGAGATTTTGGATGATGACTGT 932

RESULT 2
US-09-791-211-11
; Sequence 11, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791, 211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 11
; LENGTH: 5208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(4530)
US-09-791-211-11

Query Match 10.6%; Score 111; DB 4; Length 5208;
Best Local Similarity 54.1%; Pred. No. 2.5e-25;
Matches 272; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 396 GATACCAAGAGATGATCTGCAATAGCTTTTGGCTTGCATATTTAGTGGAGACCA 455
Db 433 GATATTGATGATGATCTGATGATGGGATGTGTGGATTTGACATGGATGGCCACCA 492
QY 456 AGTTTGAAGAAAGTGTCTCCGGGAGAGTTGCGACTGCGCATGATGTGATAGT 515
Db 493 TTATACATATGAGGGAAGT---TGCGAAGTTGCTACTAATTCAGTGTGTTCTGAG 549
QY 516 AATTATTGTGATGTATGATATTTT---CATTTGTGTATCCCTCAAGTCTCCAAAT 572
Db 550 AGCAAAATGTACTGTTCACAGTTTCTTCATGTCAGTTTTCCTCCAGGATTAAGT 609
QY 573 CTATTTGAAGATTCACATCTTTGAAGTGTGTAAGTGTGATGTGATGCTGTGAAG 632

Db 610 TTGCTTGAAGATTAAGGATTAAGGAGGAGCTGATGAGATTAAGGATTCAGTGAAA 669
QY 633 CTTTTCATGACTATGAGTATGATGATCAAGATGTGAGATCTTTGATGATTTAGCCAA 692
Db 670 CTTTACGTGACTTGTGATATCAAAATTTTGTGAGTTGACAGATGTTGCCAAT 729
QY 693 CAAAAAATGTGTGAGATTAAGAAATGGGGCTTGCTCCTACATACCTGAGACATGTTTGC 752
Db 730 AAAAAGCTGAATGTACAGAGACCTGGACCTTAAACAGTCTGGTTAAACACCTTTAGT 789
QY 753 AAAAGCTCCTGAAGCCAAACAGATCAGGCTTGGAACTGGAGATTTATCCTCTGTCA 812
Db 790 AAACAGCTCCTGAAAGACAGATCTATCCGCTGTAGCAATTTGAGATTAATTTCTCTCACT 849
QY 813 AAGCAGATTTACATACGAGCAACGATGCTTATGCTTCATGCTTACAGATTTTACAGGTT 872
Db 850 GAGACCCAGAAACGTGTGACAGCCACTGATGCTTATGCTGTTTATTTATTTACCGAAT 909
QY 873 CTTAAGGACCTTCCCTGATGCTGT 895
Db 910 TTAGAGATTTTGGATGATGACTGT 932

RESULT 3
US-09-127-670-5
; Sequence 5, Application US/09127670
; Patent No. 6228583
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute for Technology
; APPLICANT: Leonard P. Guarente
; APPLICANT: David A. Sinclair
; APPLICANT: David B. Lombard
; TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE
; FILE REFERENCE: MIT-7720DA
; CURRENT APPLICATION NUMBER: US/09/127, 670
; CURRENT FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: US 60/054, 629
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 6476
; TYPE: DNA
; ORGANISM: Murine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(4432)
US-09-127-670-5

Query Match 7.9%; Score 82.8; DB 4; Length 6476;
Best Local Similarity 53.2%; Pred. No. 3.7e-16;
Matches 199; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

QY 481 GGAAGTTGGAGCTGTCACATATGTGTAGATGATTAATTTGATGATTTGATGATATTT 540
Db 494 GCAAGATGCGAGATGATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 553
QY 541 TT---CATTTGTGTATCCCTCAAGTCTCAACATCTTATTAAGATTCACACTTGTAA 597
Db 554 CTTTCATGATGATGTTTCCCTCCAGGATTAAGAAATGTTACTAGAAACAAATCAATTAAG 613
QY 598 AGGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
Db 614 AGGACGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 673
QY 658 TCAAGATGTTGAGATCTTTACAGATTTTACCAACCAAAATTTGATGATGATGATGAT 717
Db 674 TGAAGATTTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
QY 718 GGGGCTTGGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777

Accession	Sequence	Length
Db	GGAGCCCTCAATGAGCTCTGGTTAAACACGCTATTGGGAACAACCTTTTGAAGACAAAGTCCA	793
Oy	TCAGGCTTGGGAATCTGGAGATTTTATCTCTGTCGAAGAAGCAGATTCAATFACGACGCA	837
Db	TCCCGCTGCAGCAAAATTGGAGTATATTTCCCGCTCACTGAGGACAGAAACTGATACGACCA	853
Oy	CGGATGCTTATGCT	851
Db	CTGATGCTTATGCT	867

RESULT 4

```

US-08-781-891-205
Sequence 205, Application us/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELAT
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Avenue, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
Zip: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: us/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ. ID NO.: 205:
SEQUENCE CHARACTERISTICS:
LENGTH: 4792 base pairs
type: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 145..4347
US-08-781-891-205

```

Query Match	7.78;	Score 81.2;	DB 3;	Length 4792;
-------------	-------	-------------	-------	--------------

Matches 179; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

[illegible]

QY	690	AACCAAAAATTTGGTGGAGATATAAAAAATGGGGCCCTTGCTCCACGTAACGTGACACCTTGT	749
Db	622	AATGAAAAGTTTGANGTGCGCGACACCTCGAGCCCTCAATGCTTGGTTTAAACAGCTTTA	681
QY	750	TGCNAAGAGCTCTGAAGGCCAANACAGANTCAGCCTGGGAACCTGGAGTTTATCTCTG	809
Db	682	GGGAACACACTTTTGGAAAAGACAGTCATCCGCTGACAGCANTTGGAGTAATTTCCCCCTC	741
QY	810	TCAAGCAGCAGTTACAATAAGCAGCAACGGATCTTATGCT	851
Db	742	ACTGAGGACCGAAACCTGTATGCGACCCACTGATCTTATGCT	783

RESULT 5

```

US-09-791-211-10
? Sequence 10, Application US/09791211
? Patent No. 6448080
? GENERAL INFORMATION:
? APPLICANT: Donna T. Ward
? APPLICANT: Andrew T. Walt
? TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
? FILE REFERENCE: RTS-0205
? CURRENT APPLICATION NUMBER: US/09/791,211
? CURRENT FILING DATE: 2001-02-23
? NUMBER OF SEQ ID NOS: 90
? SEQ ID NO 10
? LENGTH: 98844
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: unsure
? LOCATION: 24962
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 64383
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65468
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65469
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65470
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65471
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 87130
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 89049
? OTHER INFORMATION: unknown
? OTHER INFORMATION:
? US-09-791-211-10

```

Query Match 5.38; Score 55.4; DB 4; Length 98844;

Matches 92; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY	545	TTTGCGATCCCTCAAGCTCCACATCTATTATCAATTCACACCTTGAAAGGATG	604
Db	45421	TACAGTTTTCCCCAGGATTTAAATGTGTCTTAAATTAAGCACTTTAAAGGAGAG	45480
QY	605	TATTGGAATGATGGGACTCTGTAGACCTTTCCATACATATGGAGTTAGTATCAAGA	664
Db	45481	TGTAGGATTTGAAGGAGTACACTGAGAACTCTCAGTACTTGTATATCAATTGAAGAA	45540
QY	665	TGTTGAGATCTTTTCAGATTTATGCCAACCAAAA	697
Db	45541	TTTTTGCGAATTCACAGATGTTGCCAAATAAAAA	45573

RESULT 6

US-08-781-891-207

Sequence 207, Application US/08781891

Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

TITLE OF INVENTION: WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,891

FILING DATE: 27-DEC-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: NO. 6090620Leuberg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 207:

SEQUENCE CHARACTERISTICS:

LENGTH: 29604 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-781-891-207

Query Match 4.9%; Score 51.2; DB 3; Length 29604;

Best Local Similarity 56.5%; Pred. No. 1.5e-05;

Matches 95; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 545 TCTGTATCCCTCAAGTCTCCACATCTTATTTGAAGTTCAACACTTGTAAAGGTAGG 604

DB 19789 TTTAGTTTCCCTCCAGGAGTAAATGTTACTAGAAACAAATCAATTAAGAGCAGG 19848

QY 605 TATTTGAATGAGTGACTCTGTGAAGCTTTTCCATGACTATGAGCTTAGTTCACAAG 664

DB 19849 GGTGTGAGATTGAAGGGGACAGTGAACCTTCTGCGTATTTTGACGTCAAGTTGAGAG 19908

QY 665 TCTTTGAGATCTTTAGATTACCAACCAAAAATTTGGTGCAGATTA 712

DB 19909 TTTTGTGAGACTGACGAGATCTTCCCAATGAAAAGTAGGCGTAAATAA 19956

RESULT 7

US-08-232-463-14

Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEIFLINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: pTZgpt-Fls

US-08-232-463-14

Query Match 4.7%; Score 49.2; DB 1; Length 7218;

Best Local Similarity 11.5%; Pred. No. 2.8e-05;

Matches 33; Conservative 140; Mismatches 113; Indels 0; Gaps 0;

QY 9 ATTAATTTTATTTTCTTTCTTCTGTAAGAAATGTCATGCTCAATTTGATCGACGAC 68

DB 992 ACTATTTTCTTCTGCTTCCATACGCTCACAGAAATTAATTCGAGCTTGCTGAGGTC 1051

QY 69 GCTTTACAGAGAGAGAGCTTTCGCTATCGACGCGCATCGAAGCTTCCATATTTCTCC 128

DB 1052 GAGGAGCTTGCGATTT 1111

QY 129 CGTTCTTCTTCTTCTTCTTCTGCTGCTCGACCGCTCAAGCTCAACCTCCGTCAT 188

DB 1112 YY 1171

QY 189 GCGCAGAGAGAGAGATCAATCAATCCCAATATATATTCGTCGCAATGGCTGCTCC 248

DB 1172 YY 1231

QY 249 ATCACTTCTTCAATCTATTAACGATTTCCCTCTCTCCGTTGCC 294

DB 1232 YY 1277

RESULT 8

US-09-300-672-3

Sequence 3, Application US/09300672

Patent No. 6248937

GENERAL INFORMATION:

APPLICANT: Finkelstein, Ruth R.

APPLICANT: Lynch, Tim

APPLICANT: Goodman, Howard M.

APPLICANT: Wang, Ming-Li

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:38:19 ; Search time 27 Seconds
(without alignments)
1845.279 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491
Sequence: 1 MSSSNWIDDAFTFEELDAID.....YASWMLKYKLDLPDAVSGS 288

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_uniclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1466.5	98.4	285	10	Q9FT68
2	1439.5	96.5	313	10	Q9SVM6
3	339.5	22.8	1436	13	Q93530
4	339	22.7	968	11	Q92241
5	339	22.7	1401	11	Q9JKD4
6	330.5	22.2	643	11	Q35948
7	238	16.0	201	10	Q93VU9
8	229.5	15.4	346	5	Q9VE86
9	221.5	14.9	353	5	Q961E1
10	212.5	14.3	583	5	Q9VGN7
11	179.5	12.0	582	10	Q9C7K6
12	173.5	11.6	494	10	Q9C1V2
13	159.5	10.7	699	3	Q9C2I6
14	158	10.6	123	10	Q22964
15	151.5	10.2	199	10	Q9LKT9
16	151	10.1	496	4	Q9NVH0

17	142	9.5	599	4	Q96NP1	Q96NP1 homo sapien
18	127.5	8.6	239	10	Q9S1H3	Q9S1H3 arabidopsis
19	121	8.1	505	10	Q9FTN8	Q9FTN8 arabidopsis
20	112.5	7.5	217	10	Q80886	Q80886 arabidopsis
21	111.5	7.5	210	10	Q9LHG3	Q9LHG3 arabidopsis
22	111	7.4	242	10	Q9LHG2	Q9LHG2 arabidopsis
23	111	7.4	300	2	Q9AL74	Q9AL74 citrobacter
24	110	7.4	238	10	Q9C7A5	Q9C7A5 arabidopsis
25	105	7.0	625	5	Q9VIE1	Q9VIE1 drosophila
26	103	6.9	220	10	Q9LHG1	Q9LHG1 arabidopsis
27	102.5	6.9	710	5	Q9NA69	Q9NA69 caenorhabdi
28	101	6.8	876	5	Q17951	Q17951 caenorhabdi
29	98.5	6.6	265	10	Q9LHG5	Q9LHG5 arabidopsis
30	98.5	6.6	532	17	Q976Y3	Q976Y3 sulfolobus
31	97.5	6.5	445	17	Q9YDV0	Q9YDV0 aeropyrum P
32	97	6.5	579	2	Q9ROK2	Q9ROK2 bacillus st
33	97	6.5	906	5	Q17281	Q17281 botryllus s
34	96.5	6.5	488	16	Q97JZ0	Q97JZ0 clostridium
35	95	6.4	455	10	Q9ZWS2	Q9ZWS2 vigna mungo
36	95	6.4	925	5	Q9GUG1	Q9GUG1 caenorhabdi
37	95	6.4	4589	5	Q76506	Q76506 tetrahymena
38	94.5	6.3	353	10	Q9LHG4	Q9LHG4 arabidopsis
39	94.5	6.3	416	5	Q96144	Q96144 plasmodium
40	93	6.2	488	17	Q9HJ76	Q9HJ76 thermoplasma
41	93	6.2	552	10	Q9LVM8	Q9LVM8 arabidopsis
42	93	6.2	2224	4	Q43737	Q43737 homo sapien
43	92.5	6.2	721	10	Q94B18	Q94B18 capsicum an
44	92.5	6.2	1097	3	Q13592	Q13592 saccharomyc
45	92	6.2	237	10	Q9ZV70	Q9ZV70 arabidopsis

ALIGNMENTS

RESULT 1

ID	Q9FT68	PRELIMINARY;	PRT;	285 AA.
AC	Q9FT68;			
DT	01-MAR-2001 (TREMREL. 16, Created)			
DT	01-MAR-2001 (TREMREL. 16, Last sequence update)			
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)			
DE	EXONUCLEASE.			
GN	WRNEXO.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosidis III; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI-TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=COLUBIA; TISSUE=FLOWER;			
RA	Hartung F., Plichova H., Puchta H.;			
RT	"Molecular characterisation of RecQ homologues in Arabidopsis thaliana."			
RL	Nucleic Acids Res. 21:4275-4282(2000).			
DR	EMBL: AJ404476; CAC14871.1; -			
DR	InterPro: IPR002562; 3.5-exonuclease.			
DR	SMART: SM00474; 35EXOC; 1.			
SQ	SEQUENCE 285 AA; 31750 MW; A198CB93653E229B CRC64;			

Query Match 98.4%; Score 1466.5; DB 10; Length 285;
Best local Similarity 99.0%; Pred. No. 1.2e-126;
Matches 285; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY	1	MSSSNWIDDAFTFEELDAIDAEASYNFSSSSSSSAAPVQATTSVHGHEEDNPIN 60	
DB	1	MSSSNWIDDAFTFEELDAIDAEASYNFSSSSSSSAAPVQATTSVHGHEEDNPIN 60	
QY	61	NIRQLPRSTSTSTSYRFLPSRCRANRPFAMRGGRILYSKTEVDKRAMOLIKVLD 120	
DB	61	NIRQLPRSTSTSTSYRFLPSRCRANRPFAMRGGRILYSKTEVDKRAMOLIKVLD 120	
QY	121	KRDEGIAFVGLDIEMRPSFRKGYLPKGVATVQICVDSNYCDVMHIFRSGIPOSLOHLIE 180	

```
Db 121 KRDSGIAFVGLDIEMRSPFRKGVLPCKVATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
Qy 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSDLANOKIGDKKMWGLASTETLVCKEL 240
Db 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSDLANOKIGDKKMWGLASTETLVCKEL 240
Qy 241 LKPNRIRLGNWMEFYPLSKQOLQYATADAYASWMLYKVLKDLDPDAVSGS 288
Db 241 LKPNRIRLGNWMEFYPLSKQOLQYATADAYASWMLYKVLKDLDPDAVSGS 285

RESULT 2
Q9SYV6 PRELIMINARY; PRT; 313 AA.
ID Q9SYV6
AC Q9SYV6;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE HYPOTHETICAL 35.1 KDA PROTEIN.
GN F18A5.260 OR ATG13870.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702.
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Grueninger D., Schmidheini T., Mewes H.W., Lemcke K.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Weber N., Grueninger D., Schmidheini T., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035528; CAB36851.1; -
DR EMBL; AL161537; CAB78429.1; -
DR InterPro: IPR002562; 3.5.exonuclease.
DR Pfam: PF01612; 3.5.exonuclease; 1.
DR SMART; SM00474; 3EXOC; 1.
DR Hypothetical protein.
SQ SEQUENCE 313 AA; 35105 MW; 448D7ACC375D4D22 CRC64;

Query Match 96.5%; Score 1439.5; DB 10; Length 313;
Best Local Similarity 95.3%; Pred. No. 4e-124; 4; Indels 9; Gaps 1;
Matches 281; Conservative 1; Mismatches 281;

Qy 1 MSSSNWIDDAFTBEELLALDAIEASYNFSRSSSSSSSAAPVQATTSVHGHEEDPNQIPN 60
Db 1 MSSSNWIDDAFTBEELLALDAIEASYNFSRSSSSSSSSSAAPVQATTSVHGHEEDPNQIPN 60
Qy 61 NNRQLPRSTSTSSYKRFPLSKRANFRPMRGGRILYKATVEVDKRAMOLIKVLDLT 120
Db 61 NNRQLPRSTSTSSYKRFPLSKRANFRPMRGGRILYKATVEVDKRAMOLIKVLDLT 120
Qy 121 KRDSGIAFVGLDIEMRSPFRKGVLPCKVATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
Db 121 KRDSGIAFVGLDIEMRSPFRKGVLPCKVATVQICVDSNYCDVMHIFHSGIPQSLQHLIE 180
Qy 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSDLANOKIGDKKMWGLASTETLVCKEL 240
Db 181 DSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSDLANOKIGDKKMWGLASTETLVCKEL 240
Qy 241 LKPNRIRLGNWMEFYPLSKQOLQYATADAYASWMLYKVLKDLDPDAVSGS 286
```

```
Db 241 LKPNRIRLGNWMEFYPLSKQOLQYATADAYASWMLYKVTTKNHLITLNDLEAKIS 295

RESULT 3
ID Q93530 PRELIMINARY; PRT; 1436 AA.
AC Q93530;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DE 01-OCT-2001 (TEMBLrel. 18, Last annotation update)
DE WERNER SYNDROME HELICASE HOMOLOG (FOCUS FORMING ACTIVITY 1).
GN FFA-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus.
ON NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A. AND FUNCTION.
RX MEDLINE=98361165; Pubmed=9697700;
RA Van H., Chen C.-Y., Kobayashi R., Newport J.;
RT "replication focus-forming activity 1 and the werner syndrome gene
RL Nat. Genet. 19:375-378(1998).
RN [2]
RP FUNCTION.
RX MEDLINE=96032793; Pubmed=7569932;
RA Van H., Newport J.;
RT "FFA-1, a protein that promotes the formation of replication centers
RT within nuclei."
RL Science 269:1883-1885(1995).
CC -!- FUNCTION: ESSENTIAL FOR THE FORMATION OF DNA REPLICATION FOCAL
CC CENTERS: STABLY ASSOCIATES WITH FOCI ELEMENTS GENERATING BINDING
CC SITES FOR RP-A. EXHIBITS A MAGNESIUM DEPENDENT ATP-DEPENDENT DNA-
CC HELICASE ACTIVITY. MAY BE INVOLVED IN THE CONTROL OF GENOMIC
CC STABILITY.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RECQ SUBFAMILY OF HELICASES.
CC -!- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
DR EMBL; AF067418; AAC63512.1; -
DR InterPro: IPR002562; 3.5.exonuclease.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR002121; HRDC.
DR Pfam: PF01612; 3.5.exonuclease; 1.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00570; HRDC; 1.
DR SMART; SM00487; 3EXOC; 1.
DR SMART; SM00474; 3EXOC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00341; HRDC; 1.
DR HydroLase; Helicase; ATP-binding; DNA-binding; Nuclear protein.
KW DOMAIN
FT DOMAIN 461 467
FT DOMAIN 1098 1117 HRDC.
FT NE_BIND 512 519 ATP (BY SIMILARITY).
FT SITE 609 612 DEAD BOX.
SQ SEQUENCE 1436 AA; 161850 MW; 1BEAF05A25BAE230 CRC64;

Query Match 22.8%; Score 339.5; DB 13; Length 1436;
Best Local Similarity 35.0%; Pred. No. 4.2e-22;
Matches 82; Conservative 43; Mismatches 84; Indels 25; Gaps 7;

Qy 61 NNRQLPRSTSTSSYKRFPLSKRANFRPMRGGRILYKATVEVDKRAMOLIKVLDLT 108
Db 3 SLOKRLPEMWSVQOEDRIDDAKKSFCCKNILEDNLPFMKFNCSIYSYESNCCSLSED 62
Qy 109 KRAMOLIKVLDTRKDSGIAFVGLDIEMRSPFRKGVLPCKVATVQICVDSNYCDVMHIF- 167
Db 63 IRRSL------EEDV--LGFIDIEPPIYTKG-KTGVALIQLYCVSSKKCYLFIISP 110
Qy 168 HSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKDVEDLSDLANOKIGDKKMW 227
```

```

Db      111 MAGPPKGLKRLDESVYRKVGVEGQMKLMSDEYKLGIFLSEKANKLCKEKWT 170
QY      228 LASLTETLVCKELLPNRIRLGNMEFYPLSKOOLQVATDAYASWHLKYKVLKD 281
      171 ENGLIKRLFEOQLYKRSKSYCSNMDFLITREDQKLVATDAYAGLLIYKKLEGM 224

RESULT 4
092241
ID      092241      PRELIMINARY;      PRT;      988 AA.
AC      092241;
DT      01-MAY-1999 (Tremblrel. 10, Created)
DT      01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT      01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE      WRN PROTEIN.
GN      WRN.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      [1]
      NCBITaxID=10090;
RN      SEQUENCE FROM N.A.
RP      Paepker B.W., Gayle M., Brady M., Swartz A., Gillett L.A., Alisch R.S.,
RA      Mulligan J., Gias D., Fu Y.-H.,
RT      "Genomic structure of the human Werner's gene and cloning of the mouse
RT      homolog."
RL      Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF091216; AAC72359.1; -
DR      InterPro: IPR002562; 3.5_exonuclease.
DR      InterPro: IPR001410; DEAD.
DR      InterPro: IPR001650; Helicase_C.
DR      Pfam: PF00270; 3_5_exonuclease; 1.
DR      Pfam: PF00271; Helicase_C.
DR      Pfam: PF00474; 35EXOC; 1.
DR      SMART: SM00487; DEXDC; 1.
DR      SMART: SM00490; HELIC; 1.
KW      ATP-binding; Helicase.
SQ      SEQUENCE 988 AA; 111513 MW; DBD4D77FA55D5B2D CRC64;

Query Match      22.7%; Score 339; DB 11; Length 988;
Best Local Similarity 36.6%; Pred. No. 2,8e-22;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

QY      88 NFPMARFGRLYKTKATEVDKRAMOLIKVLDTRDESGIAFVGLDIEMRPFKGVLP 147
      37 NLPLFEPGSLIVSYEASD---CSFLSEDISMRLSDGV--VGFDMEWPPIYK----PG 86
QY      148 K---VATVOICVDSNYCDVMHIFHSIGI-POSLOHLIEDSTLVKVGIGIDGDSVKLFPHDYG 203
      87 KRSRVAVIQLCVSENKCYLFIHSSMSVFPQGLKMLENKSIKKAGVIEGQWMLLRDFD 146
QY      204 VSIKDVEDSLDANOKIGGKKMGKGLASLTETLVCKELLPNRIRLGNMEFYPLSKOOLQY 263
      147 VKLESFVELTDVANERKLCETWSLNGLVKHLGKOLLKDKSIRCSNWSNPFLETEDKLY 206
QY      264 AATDAYASWHLKYKVLKDPDPAV 285
      207 AATDAYAGLLIYQKLGNDPTV 228

RESULT 5
09JKD4
ID      09JKD4      PRELIMINARY;      PRT;      1401 AA.
AC      09JKD4;
DT      01-OCT-2000 (Tremblrel. 15, Created)
DT      01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT      01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE      WRN PROTEIN.
GN      WRN.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

```

```

OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBITaxID=10090;
      [1]
RN      SEQUENCE FROM N.A.
RP      MEDLINE-20221574; Pubmed-10757812;
RA      Lombard D.B., Beard C., Johnson B., Marciniak R.A., Dausman J.,
RA      Bronson R., Buhlmann J.E., Lipman R., Curry R., Sharpe A.,
RA      Jaenisch R., Guarente L.,
RT      "Mutations in the WRN Gene In Mice Accelerate Mortality in a p53-Null
RT      Background."
RL      Mol. Cell. Biol. 20:3286-3291(2000).
DR      EMBL; AF241636; AAF64490.1; -
DR      InterPro: IPR002562; 3_5_exonuclease.
DR      InterPro: IPR001410; DEAD.
DR      InterPro: IPR001650; Helicase_C.
DR      InterPro: IPR002121; HRDC.
DR      Pfam: PF00270; DEAD; 1.
DR      Pfam: PF00271; Helicase_C; 1.
DR      SMART: SM00474; 35EXOC; 1.
DR      SMART: SM00487; DEXDC; 1.
DR      SMART: SM00490; HELIC; 1.
DR      SMART: SM00341; HRDC; 1.
KW      ATP-binding; Helicase.
SQ      SEQUENCE 1401 AA; 157300 MW; 6CB330CD072C670 CRC64;

Query Match      22.7%; Score 339; DB 11; Length 1401;
Best Local Similarity 36.6%; Pred. No. 4.5e-22;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

QY      88 NFPMARFGRLYKTKATEVDKRAMOLIKVLDTRDESGIAFVGLDIEMRPFKGVLP 147
      37 NLPLFEPGSLIVSYEASD---CSFLSEDISMRLSDGV--VGFDMEWPPIYK----PG 86
QY      148 K---VATVOICVDSNYCDVMHIFHSIGI-POSLOHLIEDSTLVKVGIGIDGDSVKLFPHDYG 203
      87 KRSRVAVIQLCVSENKCYLFIHSSMSVFPQGLKMLENKSIKKAGVIEGQWMLLRDFD 146
QY      204 VSIKDVEDSLDANOKIGGKKMGKGLASLTETLVCKELLPNRIRLGNMEFYPLSKOOLQY 263
      147 VKLESFVELTDVANERKLCETWSLNGLVKHLGKOLLKDKSIRCSNWSNPFLETEDKLY 206
QY      264 AATDAYASWHLKYKVLKDPDPAV 285
      207 AATDAYAGLLIYQKLGNDPTV 228

RESULT 6
035948
ID      035948      PRELIMINARY;      PRT;      643 AA.
AC      035948;
DT      01-JAN-1998 (Tremblrel. 05, Created)
DT      01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT      01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE      HOMOLOG OF HUMAN WERNERS SYNDROME PROTEIN.
GN      WRN OR WRN.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBITaxID=10090;
      [1]
RN      SEQUENCE FROM N.A.
RP      Wu J., He J., Mountz J.D.,
RA      "mouse WRN."
RT      Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U97045; AAB87366.1; -
DR      MGD; MGI:109635; Wrn.
DR      InterPro: IPR002562; 3_5_exonuclease.
DR      InterPro: IPR001410; DEAD.
DR      InterPro: IPR001650; Helicase_C.
DR      Pfam: PF00271; 3_5_exonuclease; 1.
DR      Pfam: PF00271; Helicase_C; 1.
DR      SMART: SM00474; 35EXOC; 1.
DR      SMART: SM00490; HELIC; 1.

```

KW ATP-binding: Helicase.
SQ SEQUENCE 643 AA: 72821 MW: F9125C234ABB757 CRC64;
Query Match 22.2%; Score 330.5; DB 11; Length 643;
Best Local Similarity 33.3%; Pred. No. 9.4e-22;
Matches 75; Conservative 48; Mismatches 83; Indels 19; Gaps 6;
QY 61 NRRQLP-----RSINSSSTSYKREPLSRCARANFPMRFGGRILYKSTAREVDKRAQLI 115
DB 5 SLOKPEKEMWMSQORCATKEKACVQKNVLEDNLPFLFEGSIYSYEASD---CSFLS 60
QY 116 KVLDTKDESGIAFVGLDIEMRSEFKGVLPGR---VAIVQICVDSNYCDVMHIFHSGI- 171
DB 61 EDISMRLSDDY--VGFDMEMPPYIK---PGKSRVAIVQLCVSEKNCYLFHISMSVF 114
QY 172 POSLQHLIEDSVKIGIDGDSVKLFHDYGVASIKDVEDPDLANOKIGDDKKWGLASL 231
DB 115 PQGLKMLEKSKIKKAGVGEIGDQWKLRLRFDVKLESFVELIDVANKKLCACETWSLNGL 174
QY 232 TETLVCKELLKPNRIRLGMNEFYPPLSKQOLQYATADAYASWHLXK 276
DB 175 VKHVLGKQLLKDKSIRCSNMSNPFLETQKLYATADAYAGFIIVR 219
RESULT 7
ID 093VU9 PRELIMINARY; PRT; 201 AA.
AC 093VU9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE P048/H02.22 PROTEIN (P0682B08.3 PROTEIN).
GN P048/H02.22 OR P0682B08.3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_taxid=4530;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P048/H02.22";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0682B08.3";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002883; BAB67862.1; -
DR EMBL: AP003578; BAB60933.1; -
SQ SEQUENCE 201 AA: 21899 MW: 85F35C0CC5386265 CRC64;
Query Match 16.0%; Score 238; DB 10; Length 201;
Best Local Similarity 31.0%; Pred. No. 6e-14;
Matches 58; Conservative 36; Mismatches 87; Indels 6; Gaps 4;
QY 92 MREGGRILYKSTAREVDKRAQLIYDTRKDESGIAFVGLDIEMRSEFKGVLPGRVAT 151
DB 10 VREGSATIDITTVISDVAADEMARGVAAARAGRGRL-IVGLDCWKPN-HVSWKTSKVAV 67
QY 152 VOICVDSNYCDVMHIFHSGI-IPSLQHLIEDSVKIGIDGDSVKLFHDYGVASIKDVE 210
DB 68 LQLCAGRREFLVQLDFRANVPPRAVDLDDPSRLVGLTGVEBAALAEADYGCWCAAPV 127
QY 211 DLSDLANQK---IGDDKKWGLASLTTLVCKELLKPNRIRLGMNEFYPPLSKQOLQYATD 267
DB 128 DLSDLANQK---IGDDKKWGLASLTTLVCKELLKPNRIRLGMNEFYPPLSKQOLQYATD 267
QY 268 AVASWHL 274

DB 188 AVASWHL 194
RESULT 8
ID 09VE86 PRELIMINARY; PRT; 346 AA.
AC 09VE86;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CG7670 PROTEIN.
GN CG7670.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllozoa; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jhalali M., Kalush F., Kapen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheetler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003721; AAP55541.1; -
DR FlyBase: FBgn0038608; CG7670.
DR InterPro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
SQ SEQUENCE 346 AA: 39398 MW: 3150A9B4EBF74EB0 CRC64;
Query Match 15.4%; Score 229.5; DB 5; Length 346;
Best Local Similarity 26.9%; Pred. No. 7.7e-13;
Matches 66; Conservative 43; Mismatches 123; Indels 13; Gaps 5;
QY 41 TVQATTSVHGHEDPDNPINIRKRLPRSTTSSTSYKRFPLSKCARANFPMRFGGRILY 100

Db	68	TPDVTEKRLAMEBENP----	PKRRSSLUTSTRSTBMA	DGSPSPKEKPEKLP	IKYGAIKY	124
Qy	101	SKTATEVDRKAMOLIKVLDTRKDESGIA	FVGLDIDMRSPFRKKVGLPGK	YATVQICVDSNY	160	
Db	125	FTESODDIASADVDVQWWEKQAE--	VVPMAFDMWEPSPFQFG--	PGKSAV	180	
Qy	161	CDVNHIFH-SGIPOSLOH	IEDSTLVKIGIGIDGSVLFHDX-----	GVSIKDVEDLSD	214	
Db	181	CYIQTIVTWKRLPALVALINHPKRLHGVN	IKNDPFRKLARPEPVTAEPLEK	CVDLGL	240	
Qy	215	LANOKIGDKRWGLASLTETLVCKELLP	PNRIIRLGNMEFEYPLSKOOL	YATADTAYASMHL	274	
Db	241	MCNEVCETGCRWMSLERLNTFLNFTAK	KAMDSKKVRMSKMHIYIPLEDENQ	LMTAIDYIGVYI	300	
Qy	275	YKVLK	279			
Db	301	YRELE	305			
RESULT 9						
ID	0961E1	PRELIMINARY;	PRT;	353	AA.	
AC	0961E1					
DT	01-DEC-2001 (TREMBLrel. 19,	Created)				
DT	01-DEC-2001 (TREMBLrel. 19,	Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19,	Last annotation update)				
DE	GM4514P.					
GN	CG670.					
OS	Drosophila melanogaster (Fruit fly).					
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
OC	Ephyridiida; Drosophilidae; Drosophila.					
OX	NCBI_TaxID=7227;					
RP	SEQUENCE FROM N.A.					
RA	Stapleton M., Brockstein P., Hong L., Aghayani A., Carlson J.,					
RA	Champe M., Chavez C., Dorsett V., Fattah D., Frise E., George R.,					
RA	Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,					
RA	Nunoo J., Paolel J., Paragas V., Park S., Phouanenvong S., Wan K.,					
RA	Yu C., Lewis S.E., Rubin G.M., Celniker S.,					
RL	Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.					
DR	EMBL: AY051647; AAK93071.1; -					
SO	SEQUENCE 353 AA; 40248 MW; F9836CB263868AB4 CRC64;					
Query Match 14.9%; Score 221.5; DB 5; Length 353;						
Best Local Similarity 26.5%; Pred. No. 4,3e-12;						
Matches 65; Conservative 43; Mismatches 124; Indels 13; Gaps 5;						
Qy	41	TVQATTSVHGHEEDPNQIPNNIRROL	PPSITSSISYKRPPLSRCARNEP	AMREGGRILY	100	
Db	75	TPDVTEKRLAMEBENP----	PKRRSSLUTSTRSTBMA	DGSPSPKEKPEKLP	IKYGAIKY	131
Qy	101	SKTATEVDRKAMOLIKVLDTRKDESGIA	FVGLDIDMRSPFRKKVGLPGK	YATVQICVDSNY	160	
Db	132	FTESODDIASADVDVQWWEKQAE--	VVPMAFDMWEPSPFQFG--	PGKSAV	180	
Qy	161	CDVNHIFH-SGIPOSLOH	IEDSTLVKIGIGIDGSVLFHDX-----	GVSIKDVEDLSD	214	
Db	188	CYIQTIVTWKRLPALVALINHPKRLHGVN	IKNDPFRKLARPEPVTAEPLEK	CVDLGL	247	
Qy	215	LANOKIGDKRWGLASLTETLVCKELLP	PNRIIRLGNMEFEYPLSKOOL	YATADTAYASMHL	274	
Db	248	MCNEVCETGCRWMSLERLNTFLNFTAK	KAMDSKKVRMSKMHIYIPLEDENQ	LMTAIDYIGVYI	307	
Qy	275	YKVLK	279			
Db	308	YRELE	312			
RESULT 10						
ID	09YGN7	PRELIMINARY;	PRT;	583	AA.	
AC	09YGN7;					

01-MAY-2000 (Tremblrel. 13, Created)
01-MAR-2001 (Tremblrel. 16, Last sequence update)
01-DEC-2001 (Tremblrel. 19, Last annotation update)
CG6744 PROTEIN.
CG6744.
GN
CN
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:
OC Ephydroidae: Drosophilidae: Drosophila.
OX NCBI_taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RC MEDLINE=20196006; Pubmed=10731132.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chape M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Klupp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei Y., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Paclik J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Stiden-Klamos I., Stimpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskiy R., Tecor C., Turner R., Venter E., Weiss A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodge T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003691; AAF54639.2;
DR FLYBase; FBgn0037901; CG6744.
DR InterPro; IPR002562; 3_5_exonuclease.
DR SMART; SM00474; 35EXOC; 1.
SQ SEQUENCE 583 AA; 66866 MW; 562984B17795ZAD0 CRC64;

```

OY 280 DL 281
DB 227 DL 228

RESULT 11
O9C7K6 PRELIMINARY; PRT; 582 AA.
AC O9C7K6:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TEMBLrel. 18, Last annotation update)
DE 3'-5' EXONUCLEASE, PUTATIVE.
GN F14G9.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucrids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maill R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Soultwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utechtback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
DR EMBL: AC069159; AAG50917.1;
DR InterPro: IPR002562; 3_5-exonuclease.
DR Pfam: PF01612; 3_5-exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
KW Exonuclease.
SQ SEQUENCE 582 AA; 66467 MW; 34D752473E786680 CRC64;

Query Match 12.0%; Score 179.5; DB 10; Length 582;
Best Local Similarity 28.2%; Pred. No. 6.2e-08;
Matches 53; Conservative 36; Mismatches 66; Indels 33; Gaps 7;

OY 126 GIAFVGLDIEMRSFKGVLPKAVATVQICVDSN--YCDVMHIFHSG--IPOSLOHLIE 180
DB 377 GCRVGVIDCEWKNYIKSKONKVSIMQIGSDTKIFILDIKILYNASELIDNCILSHIQ 436
OY 181 DSTLVFY-----GIGIDGSVKLFHDYGSIKDVE-----DLSDLNOKI 220
DB 437 SKETLVLVSTEDYPPHKASSGYNFOCDIKQLALSTG-DKCEERYDMLIDIONVNEEP 495
OY 221 GGDKKMGSLATETLVCKELKPNRIRLGNMEFYPLSKOOLQYAADAYASMLLYKLD 280
DB 496 G-----GLAGLTGKILIGVSLNKR--RNSDMWRPLSQNLERAAALDAANLVLIHFNRVD 548
OY 281 LPDAVSGS 288
DB 549 HPHDSSS 556

RESULT 12
O93VS2A

```

```

ID 093VS2 PRELIMINARY; PRT; 494 AA.
AC 093VS2:
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE P0487H02.21 PROTEIN (P0682B08.2 PROTEIN).
GN P0487H02.21 OR P0682B08.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0487H02."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0682B08."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002883; BAB67861.1;
DR EMBL: AP003578; BAB60932.1;
SQ SEQUENCE 494 AA; 51497 MW; 8F83F79405C2BD1E CRC64;

Query Match 11.6%; Score 173.5; DB 10; Length 494;
Best Local Similarity 24.3%; Pred. No. 1.8e-07;
Matches 69; Conservative 42; Mismatches 126; Indels 47; Gaps 11;

OY 13 EEBLLAIDAEAYNRSRSSSSSSAAPTQATSVHG-----HEEDN 56
DB 204 DEELMG---SAGRRARPPSSSMGVGEATEGTGIGGPPATSSSSPPPPHPSH 259
OY 57 QINNNIRPOLRSITSTSYKRRPLSRCRANRPNRPFGRILYSTATVEVKRAMQLIK 116
DB 260 HTHFFDRSIDRSA-----PLLDVDGGCTVFSFSAI-DTIVSDDAAADAEWR 309
OY 117 -----VLDTRKDESGIAFGLDIEMRP--SEKGVLPKAVATVQICVDSNYSYCDVMHIFHS 169
DB 310 RVRASATTPRGCGGL-LVGLDCEMKRCDHLPAPVAP-TVAILQLCAGS-CLIQDLHY 366
OY 170 G-----IPOSLOHLIEPSTLVKVGIGIDGSVKLFHDYGSIKVEPLSLANOKIG---G 222
DB 367 AGARRVPPLVGLDLADPSVRLVGLIGENNAKLADGIVGCAAPVLEDCVDRRLRLPG 426
OY 223 DKRWGLASLETETLVCKELKPNRIRLGNMEFYPLSKOOLQYANT 266
DB 427 ARRLGLKGYREVLTGLTMEKPMQVTRSDWERRRLDAAGVRYACS 470

RESULT 13
O9C2I6 PRELIMINARY; PRT; 699 AA.
AC 09C2I6:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE RELATED TO WERNER SYNDROME HELICASE.
GN 93G11.60.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN (1)
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:38:04 : Search time 13 Seconds
(without alignments)
857.787 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491
Sequence: 1 MSSSNMIDDAFTEELALD.....YASMLYKVLKDLPAVSGS 288

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	352	23.6	1432	1 WRN_HUMAN	014191 homo sapien
2	339	22.7	1401	1 WRN_MOUSE	009053 mus musculu
3	136.5	9.2	910	1 Y068_CAEEL	P34607 caenorhabdi
4	95	6.4	508	1 CPVL_RAT	P22443 rattus norv
5	93.5	6.3	1525	1 POLG_STEVM	P09732 s genome po
6	93	6.2	2224	1 FAS_HUMAN	P12259 homo sapien
7	93	6.2	3079	1 IRA2_YEAST	P19158 saccharomyc
8	92	6.2	758	1 LEU2_SCHPO	014289 schizosacch
9	91.5	6.1	473	1 VL2_HPV03	P36744 human papil
10	91	6.1	503	1 DPO1_MOUSE	P28649 mus musculu
11	89.5	6.0	1324	1 CPV1_HAEIN	P43741 haemophilus
12	89.5	6.0	1487	1 BLM_DROME	004716 arabidopsis
13	89.5	6.0	1487	1 BLM_DROME	004716 arabidopsis
14	88	5.9	259	1 HK22_BRARE	090461 brachydanio
15	88	5.9	416	1 PROA_VIBCH	09XPE9 vibrio chol
16	88	5.9	784	1 Y063_CAEEL	P34603 caenorhabdi
17	87	5.8	579	1 MTC1_BACST	P43423 bacillus st
18	87	5.8	3137	1 CA36_CHICK	P15989 gallus gall
19	86.5	5.8	346	1 PST5_ECOLI	036128 escherichia
20	86.5	5.8	470	1 PLSR_CUCSA	039639 cucumis sat
21	86	5.8	552	1 FMR2_APLCA	P08021 aplysia cal
22	85.5	5.7	1220	1 DEOC_MYCPI	P47722 mycoplasma
23	85.5	5.7	1394	1 E75B_DROME	P17672 drosophila
24	85.5	5.7	2298	1 C215_HUMAN	09Y315 homo sapien
25	84.5	5.7	480	1 UGDH_SOYBN	096558 glycine max
26	84.5	5.7	830	1 GYRA_CLOAB	P94605 clostridium
27	83.5	5.6	777	1 YASB_SCHPO	010146 schizosacch
28	83	5.6	482	1 6PGD_HUMAN	P52209 homo sapien
29	83	5.6	1129	1 PHVB_SOLTU	P34094 solanum tub
30	82	5.5	229	1 RPE_CHLPN	092829 chlamydia p
31	82	5.5	492	1 6PGD_SCHPO	P78812 schizosacch
32	82	5.5	540	1 CH60_THERR	060024 thermoaer
33	81.5	5.5	636	1 GYRB_THDMA	P77993 thermotoga

34	81.5	5.5	1002	1 YEMA_DROME	P25992 drosophila
35	81	5.4	309	1 CC2B_ARATH	P25859 arabidopsis
36	81	5.4	459	1 Y819_PYRHO	058549 pyrococcus
37	81	5.4	560	1 VAOX_PENSI	P56216 penicillium
38	81	5.4	864	1 CHEA_BORBU	044737 borrelia bu
39	81	5.4	885	1 PNC2_HUMAN	001780 homo sapien
40	80.5	5.4	399	1 RRP6_P14HA	P22044 human parat
41	80.5	5.4	451	1 NFS1_MOUSE	092133 mus musculu
42	80.5	5.4	473	1 VL2_HPV28	P50799 human papil
43	80.5	5.4	543	1 PROA_LEGPN	P21347 legionella
44	80.5	5.4	733	1 RRP6_YEAST	012149 saccharomyc
45	80	5.4	574	1 DPO1_AQUAE	067779 aquifex aeo

ALIGNMENTS

RESULT 1
WRN_HUMAN STANDARD; PRT; 1432 AA.
AC Q14191;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Werner syndrome helicase.
GN WRN OR RECQL2 OR RECQ3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96181115; PubMed=8602509;
RA Yu C.-E., Oshima J., Fu Y.-H., Wajsmann E.M., Hisama F., Alisch R.,
RA Matthews S., Nakura J., Miki T., Quais S., Martin G.M., Mulligan J.,
RA Scheinberg G.D.;
RT "Positional cloning of the Werner's syndrome gene";
RL Science 272:258-262(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Peepker B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA Mulligan J., Galas D., Fu Y.-H.;
RT "Genomic structure of the human Werner's gene and cloning of the
RT mouse homolog.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=98284027; PubMed=9618508;
RA Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.;
RT "Nucleolar localization of the Werner syndrome protein in human
RT cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).
RN [4]
RP REPEATS.
RX MEDLINE=99160561; PubMed=10049920;
RA Kusano K., Berres M.E., Engels W.R.;
RT "Evolution of the RECQ family of helicases: A drosophila homolog,
RT Dmblm, is similar to the human Bloom syndrome gene.";
RL Genetics 151:1027-1039(1999).
RN [5]
RP REVIEW ON VARIANTS.
RX MEDLINE=99235545; PubMed=10220139;
RA Moser M.J., Oshima J., Monnat R.J., Jr.;
RT "WRN mutations in Werner syndrome.";
RL Hum. Mutat. 13:271-279(1999).
RN [6]
RP VARIANT ARG-1367.
RX MEDLINE=9713161; PubMed=9021029;
RA Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakuugi H.,
RA Ikegami H., Hiyaki J., Edland S.D., Martin G.M., Ogihara T.;
RT "Association of a polymorphic variant of the Werner helicase gene with
RT myocardial infarction in a Japanese population.";
RL Am. J. Med. Genet. 68:494-498(1997).

RN [7]
 RP ERRATUM.
 RA Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
 RA Ikegami H., Higaki J., Egland S.D., Martin G.M., Oghihara T.,
 RL Am. J. Med. Genet. 70:103-103(1997).
 RN [8]
 RP VARIANTS ILE-387 AND LEU-1074.
 RX MEDLINE=98111850; PubMed=9450180;
 RA Messelitzer C., Rupplitsch W., Weirich-Schwaiger H., Weirich H.G.,
 RA Jabkowsky J., Klein G., Schweizer M., Hirsch-Kauffmann M.,
 RT "Werner syndrome: characterization of mutations in the WRN gene in an
 RT affected family.";
 RL Eur. J. Hum. Genet. 5:364-370(1997).
 RN [9]
 RP VARIANT ILE-387.
 RA Vidal V., Bay J.-O., Champomier F., Grancho M., Beauville L.,
 RA Glowaczower C., Lemery D., Ferreira M., Bignon Y.-J.;
 RT "The 1396del A mutation and a missense mutation or a rare polymorphism
 RT of the WRN gene detected in a French Werner family with a severe
 RT phenotype and a case of an unusual vulvar cancer.";
 RL Hum. Mutat. 11:413-414(1998).
 RN [10]
 RP VARIANTS ALA-324 AND ARG-1367.
 RX MEDLINE=99167244; PubMed=10069711;
 RA Casto E., Ogburn C.E., Hunt K.E., Tlavis R., Louhija J.,
 RA Penttinen R., Erkkola R., Panduro A., Riestra R., Plussan C.,
 RA Deeb S.S., Wang L., Edland S.D., Martin G.M., Oshima J.;
 RT "Polymorphisms at the Werner locus: I. Newly identified polymorphisms,
 RT ethnic variability of 1367C>G/Arg, and its stability in a population
 RT of Finnish centenarians.";
 RL Am. J. Med. Genet. 82:399-403(1999).
 CC -1- FUNCTION: ESSENTIAL FOR THE FORMATION OF DNA REPLICATION FOCAL
 CC CENTERS; STABLY ASSOCIATES WITH FOCI ELEMENTS GENERATING BINDING
 CC SITES FOR RP-A. EXHIBITS A MAGNESIUM-DEPENDENT ATP-DEPENDENT DNA-
 CC HELICASE ACTIVITY. MAY BE INVOLVED IN THE CONTROL OF GENOMIC
 CC STABILITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
 CC -1- DISEASE: DEFECTS IN WRN ARE THE CAUSE OF WERNER SYNDROME (WS): A
 CC RARE AUTOSOMAL RECESSIVE PROGEROID SYNDROME CHARACTERIZED BY THE
 CC PREMATURE ONSET OF MULTIPLE AGE-RELATED DISORDERS, INCLUDING
 CC ATHEROSCLEROSIS, CANCER, NON-INSULIN-DEPENDENT DIABETES MELLITUS
 CC (NIDDM), OCULAR CATARACTS AND OSTEOPOROSIS. THE MAJOR CAUSE OF
 CC DEATH (AT A MEDIAN AGE OF 47) IS MYOCARDIAL INFARCTION (MI).
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECO SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
 CC -1- DATABASE: NAME=WRN; NOTE=WRN mutation db (Werner disease);
 CC WWW="http://www.pathology.washington.edu/werner/ws_wrn.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; L76937; AAC41981.1; -;
 DR EMBL; AF091214; AAC63361.1; -;
 DR EMBL; AF181897; AAF06162.1; -;
 DR EMBL; AF181896; AAF06162.1; JOINED.
 DR MIM; 604611; -;
 DR MIM; 277700; -;
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002121; HRDC.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam; PF01612; 3_5_exonuclease; 1.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF00570; HRDC; 1.
 DR SMART; SM00474; 3SEXC; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICC; 1.

DR SMART; SM00341; HRDC; 1.
 KW Hydrolyase; Helicase; ATP-binding; DNA-binding; Nuclear protein;
 KW Polymorphism; Repeat.
 FT DOMAIN 424 477
 FT
 FT
 FT REPEAT 424 450
 FT REPEAT 451 477
 FT DOMAIN 507 510
 FT DOMAIN 1150 1229
 FT NP_BIND 571 578
 FT SITE 668 671
 FT VARIANT 324 324
 FT
 FT VARIANT 387 387
 FT
 FT VARIANT 1074 1074
 FT
 FT VARIANT 1367 1367
 FT
 FT
 FT SEQUENCE 1432 AA; 162494 MW; DE02C0059F7B62EB CRC64;
 SQ
 Query Match 23.6%; Score 352; DB 1; Length 1432;
 Best Local Similarity 37.7%; Pred. No. 8, 8e-22;
 Matches 75; Conservative 41; Mismatches 75; Indels 8; Gaps 3;
 QY 88 NPFAMRPGGRILVSKATREDKRAMQLIKVLDKRDSEGIAPVGLDIEMRSPRKGLPG 147
 DB 43 DLPLEFTGSTIVSYNDSQSFSE-----DISMSISDGDVGFPMENPPLYNRKLLG 95
 QY 148 KVATVQICVDSNYCDVMHIFHSGI-POSTQHLIEDSTLVKVGIGIDGSKLFHDYGVSI 206
 DB 96 KVALIDQVSEKSCYLFHVSMSGVFPOGLKMLENKAVKAGVIGDQWKLRLDPDIKL 155
 QY 207 KVEDLSLANQKIGGDKKGLASLFTLYCKELLPNRLRLGNMEFFPLSKOLOYAAT 266
 DB 156 KNFVELTDVANKRLKCTEWSLSLVKHLKQLDLKDRCSIRCSNWSKFLTEQOKLYAAT 215
 QY 267 DAYASMLYKLVKLDLPAV 285
 DB 216 DAYAGFTIYNLEIIDTV 234
 RESULT 2
 WRN_MOUSE STANDARD; PRT; 1401 AA.
 ID WRN_MOUSE 009053; 009050; 092242;
 AC 009053; 009050; 092242;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Werner syndrome helicase homolog.
 GN WRN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C; TISSUE=Testis, and Spleen;
 RX MEDLINE=97288537; PubMed=9143515;
 RA Imanura O., Ichikawa K., Yamabe Y., Goto M., Sugawara M.,
 RA Furuchi Y.;
 RT "Cloning of a mouse homologue of the human Werner syndrome gene and
 RT assignment to 8A4 by fluorescence in situ hybridization.";
 RL Genomics 41:298-300(1997).
 RN [2]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=98284027; PubMed=9618508;
 RA Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.;
 RT "Nucleolar localization of the Werner syndrome protein in human
 RT cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).

```

RN [3]
SEQUENCE FROM N.A.
RA Peepers B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA Mulligan J., Galas D., Fu Y.-H.;
RT "Genomic structure of the human Werner's gene and cloning of its mouse
RT homolog.";
CC Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF GENOMIC STABILITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECD SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D86527; BAA20270.1; -
DR EMBL: D86526; BAA20269.1; -
DR EMBL: AF091215; AAC78077.1; -
DR MGD: MGI:109635; Wtn
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001211; HRDC.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00370; HRDC; 1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR SMART: SM00341; HRDC; 1.
KM Helicase; ATP-binding; Nuclear protein.
FT NP_BIND 535 542 ATP (BT SIMILARITY).
FT SITE 632 635 DEAD BOX.
FT DOMAIN 1115 1194 POLY-SER.
FT DOMAIN 1387 1390 HRDC.
FT CONFLICT 101 101 N -> S (IN REF. 3).
FT CONFLICT 228 228 V -> A (IN REF. 3).
FT CONFLICT 250 250 L -> S (IN REF. 3).
FT CONFLICT 452 452 M -> V (IN REF. 3).
FT CONFLICT 459 459 K -> T (IN REF. 3).
FT CONFLICT 468 468 C -> R (IN REF. 3).
FT CONFLICT 619 619 K -> Q (IN REF. 3).
FT CONFLICT 800 800 Q -> K (IN REF. 3).
FT CONFLICT 1021 1021 L -> S (IN REF. 3).
FT CONFLICT 1145 1145 A -> T (IN REF. 3).
FT CONFLICT 1181 1182 VG -> LE (IN REF. 3).
FT CONFLICT 1252 1252 V -> A (IN REF. 3).
FT CONFLICT 1308 1308 I -> L (IN REF. 3).
FT CONFLICT 1356 1356 V -> A (IN REF. 3).
SQ SEQUENCE 1401 AA; 157256 MW; 94906092467FB8C CRC64;

Query Match 22.7%; Score 339; DB 1; Length 1401;
Best Local Similarity 36.6%; Pred. No. 1.1e-20;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

```

```

OY 264 AATATYASWHLKYKLPDPAY 285
Db 207 AATDAYAGLIYOKLGNIGDTV 228

RESULT 3
Y068_CAEEL STANDARD; PRT; 910 AA.
AC P34607;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 105.6 kDa protein ZK1098.8 in chromosome III.
GN ZK1098.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; Pubmed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirsten J., Laister N.,
RA Lettelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaildon C., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Watson L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
KM Mature 368,32-38(1994).
CC -!- SIMILARITY: TO RIBONUCLEASE D.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Z22176; CAA80137.1; -
DR PTR: S40930; S40930.
DR WormPep: ZK1098.8; CE00370.
DR InterPro: IPR002562; 3_5_exonuclease.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR SMART: SM00474; 35EXOC; 1.
KM Hypothetical protein.
SQ SEQUENCE 910 AA; 105569 MW; 5512D15423517FCD CRC64;

Query Match 9.2%; Score 136.5; DB 1; Length 910;
Best Local Similarity 24.8%; Pred. No. 0.00094;
Matches 77; Conservative 34; Mismatches 110; Indels 89; Gaps 13;

```



```

FT INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
FT CHAIN 1 121 CELLULAR AMINOPEPTIDASE.
FT PROPEP 122 213 CAPSID PROTEIN C.
FT CHAIN 214 288 ENVELOPE GLYCOPROTEIN M.
FT CHAIN 289 789 MAJOR ENVELOPE PROTEIN E.
FT CHAIN 790 1203 NONSTRUCTURAL PROTEIN NS1.
FT CHAIN 1204 1368 NONSTRUCTURAL PROTEIN NS2A.
FT CHAIN 1369 1499 NONSTRUCTURAL PROTEIN NS2B.
FT CHAIN 1500 >1525 PROTEASE/HELICASE (NS3).
FT TRANSMEM 108 119 POTENTIAL.
FT TRANSMEM 253 268 POTENTIAL.
FT TRANSMEM 274 288 POTENTIAL.
FT TRANSMEM 751 762 POTENTIAL.
FT TRANSMEM 768 787 POTENTIAL.
FT TRANSMEM 1173 1188 POTENTIAL.
FT DISULFID 291 318 BY SIMILARITY.
FT DISULFID 348 404 BY SIMILARITY.
FT DISULFID 362 393 BY SIMILARITY.
FT DISULFID 380 409 BY SIMILARITY.
FT DISULFID 478 576 BY SIMILARITY.
FT DISULFID 593 624 BY SIMILARITY.
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 996 996 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1189 1189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 1525 1525 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 1525 AA; 167891 MW; ELA373FLE511159F CRC64;

Query Match
Best Local Similarity 6.3%; Score 93.5; DB 1; Length 1525;
Matches 47; Conservative 34; Mismatches 85; Indels 45; Gaps 11;

QY 82 SRRANFPMRGRGRLISKATFVDKRAMQLIKVLDTRKDE---SGIAFYGLDIE-WR 137
DB 764 ARDRSISLTLAVGILIFLATVQADSGC---AISLRRLKCGGGIFVNDVEKWK 818
QY 138 PSFRKGVL--PGKVATVOICVDNVCVDM-----HIFHGIFOSLOHLEDSTLVVGI 189
DB 819 SDKKRPLPTPTGLAHYIQEHANGYCGIRSTRLHLMENIOLRELAIFEDNE-IDLSV 877
QY 190 GIDGDSVKLFHDYGVSIKDEVDLSLANOKIGDKKMGASLTETLVCKELLPNRIILG 249
DB 878 VVOGEDP-KYYKRAPRRRLKLEDELDY-----GMKKWKGTLFEVPE-----RLG 918
QY 250 NMEEFY---PLSKQQLQYAAATDAASWHLRYV 277
DB 919 NMTEVVVDGPETKE-----CPTANRANNSFKV 944

RESULT 6
FA5_HUMAN STANDARD: PRT; 2224 AA.
AC P12259; Q14285;
DC 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coagulation factor V precursor (Activated protein C cofactor).
GN F5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBL_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9223268; PubMed=1567832;
RA Cripe L.D., Moore K.D., Kane W.H.;
RT "Structure of the gene for human coagulation factor V.";
RN [2] Biochemistry 31:377-3785(1992).

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=87260886; PubMed=3110773;
RA Jenny R.J., Pittman D.D., Toole J.J., Kriz R.W., Aldape R.A.,
RA Hewick R.M., Kaufman R.J., Mann K.G.;
RT "Complete cDNA and derived amino acid sequence of human factor V.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
RN [3]
RP SEQUENCE OF 1-1600 FROM N.A.
RX MEDLINE=88107560; PubMed=2827731;
RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;
RT "Cloning of cDNAs coding for the heavy chain region and connecting
RT region of human factor V, a blood coagulation factor with four types
RL of internal repeats.";
RN [4]
RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.
RX MEDLINE=86313665; PubMed=3092220;
RA Kane W.H., Davie E.W.;
RT "Cloning of a cDNA coding for human factor V, a blood coagulation
RT factor homologous to factor VIII and ceruloplasmin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
RN [5]
RP PARTIAL SEQUENCE FROM N.A.
RC TISSUE-Fibroblast;
RX MEDLINE=93203619; PubMed=8454869;
RA Shen N.L., Fan S.-T., Pyatt J., Graff R., Lapolla R.J.,
RA Edington T.S.;
RT "The serine protease cofactor factor V is synthesized by
RT lymphocytes.";
RL J. Immunol. 150:2992-3001(1993).
RN [6]
RP SULFATION.
RX MEDLINE=94264012; PubMed=8204629;
RA Pittman D.D., Tomkinson K.N., Michnick D., Seligsohn U.,
RA Kaufman R.J.;
RT "Posttranslational sulfation of factor V is required for efficient
RT thrombin cleavage and activation and for full procoagulant activity.";
RL Biochemistry 33:6952-6959(1994).
RN [7]
RP SULFATION.
RX MEDLINE=9036699; PubMed=2168225;
RA Horton G.L.;
RT "Sulfation of tyrosine residues in coagulation factor V.";
RL Blood 76:946-952(1990).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2065-2224.
RX MEDLINE=20052169; PubMed=10586886;
RA Macedo-Ribeiro S., Bode W., Huber R., Quinn-Alten M.A., Kim S.W.,
RA Ortel T.L., Bourenkov G.P., Bartunik H.D., Stubbs M.T., Kane W.H.,
RA Fuentes-Prior P.;
RT "Crystal structures of the membrane-binding C2 domain of human
RT coagulation factor V.";
RL Nature 402:434-439(1999).
RN [9]
RP VARIANT APCR G1N-534.
RX MEDLINE=94217810; PubMed=8164741;
RA Bertina R.M., Koelseman B.P.C., Koster T., Rosendaal F.R.,
RA Dirven R.J., de Ronde H., van der Velden P.A., Reitsma P.H.;
RT "Mutation in blood coagulation factor V associated with resistance to
RT activated protein C.";
RL Nature 369:64-67(1994).
RN [10]
RP FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.
CC - SUBUNIT: Factor Va is composed of a heavy chain and a light
CC chain, noncovalently bound. The interaction between the two chains
CC is calcium-dependent.
CC - DOMAIN: DOMAIN B CONTAINS 35 X 9 AA TANDEN REPEATS, AND 2 X 17 AA
CC REPEATS.
CC - PTM: Thrombin activates factor V proteolytically to the active
CC cofactor, factor Va (formation of a heavy chain at the N-
CC terminus and a light chain at the C-terminus).
CC - PTM: SULFATION IS REQUIRED FOR EFFICIENT THROMBIN CLEAVAGE AND
CC ACTIVATION AND FOR FULL PROCOAGULANT ACTIVITY.

```


Query Match 6.2%; Score 92; DB 1; Length 758;
 Best Local Similarity 18.2%; Pred. No. 4.4;
 Matches 61; Conservative 53; Mismatches 107; Indels 114; Gaps 17;

QY 34 SSSSAAPTQATTS-----VHCEDDPNOIPNNIR-ROLPRS-----ITSSTSYKREPL 81
 DB 292 NADAVLPTVWGTSPDVIPIINGNIPDPAHVKNVRAASIORSELYGKLPNTSIVSYPI 351
 QY 82 SR-----CRARNFPAMRFGGRILYKSTATVEVDRAM-----OLIKYDTRDSESGI--AF 129
 DB 352 DKAFITGSCNSTRIEDRLAAAVKGRVAVANVADAMIVPGSLVKW--AAEAGLDQIF 408
 QY 130 VGDIEMRP-----SFRKGVLPKGVATVOIC-----V 156
 DB 409 IEAGFDMREAGCSMCLGMNDOLKPYERCASSTNRNPECGQAGKGRTHLYSPAMAAAI 468
 QY 157 DSNYCDVMHIF---HSGIPDLOHLIEDSTLVVVGIGIDGDSKYLPHDVGVSIRKVEDL- 212
 DB 469 KGHLCNVREFGVDVSNGPSPIITNKVDPSS-----HDVEG-----DGLSYDADTADAV 516
 QY 213 --SDLANOKIGDPKWKGGLAFL-----TETLVCKELKLP-NRIIRLN 250
 DB 517 TDADGATNAGSVSSGSGACIPKFTYVEGIAAPLPMANVTDXIIRKQFLTKRTGLGO 576

QY 251 WEFPYPLSKOQLOVAAVDAYASWHLKYVKLDLPDPAV 285
 DB 577 FAFY-----EIRYDADG-----KEIPDFV 595

RESULT 9
 ID VLD_HPV03 STANDARD; PRT; 473 AA.
 AC P36744;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Minor capsid protein L2.
 GN L2.
 OS Human papillomavirus type 3.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 RN NCBI_TaxID=10614;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94265501; PubMed=8205838;
 RX Delius H., Hofmann B.;
 RT "Primer-directed sequencing of human papillomavirus types."
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X74462; CAA52473.1; -
 DR PIR: S36553;
 DR InterPro: IPR000784; Late_L2.
 DR Pfam: PF00513; Late_Protein_L2; 1.
 KW Coat protein; Late protein.
 SQ SEQUENCE 473 AA; 50662 MW; DD8298FE517A04DD CRC64;

Query Match 6.1%; Score 91.5; DB 1; Length 473;
 Best Local Similarity 21.6%; Pred. No. 2.6;
 Matches 66; Conservative 34; Mismatches 95; Indels 111; Gaps 14;

QY 1 MSSSMNIDAFTEELALDAIEASYNFSRSSSSSAAPTQATTSVHGEDDPNOIPN 60
 DB 153 VSTNSNPAFTPEPLLEV-----QNEVSGHILISFTSGTHYEE----- 195
 QY 61 NTRQLPRSTSTSYKRPPLSRARNPAM-RFGGRILYKSTATVEVDRAMOLIKVLD 119

DB 196 -----IPMERFASPCGTETPIS---STPPVGSRIAGPRILYSAVAYQY-----KVTD 239
 QY 120 ---TKRDESGIAFVGLDIEMRPS---FRKGLPGKATVQICVDSNYCDVMHIF----- 167
 DB 240 PAFLTRPRRSIMTDPNPFEPEDETIIFERYSPSOVP-----DSDFLDIRLRHAPLNS 293
 QY 168 -----HSGIPDLOHLIEDSTLVKVGIGIDGDSVKLFHDYGVSIKDVEDLSDLANOKIG 221
 DB 294 RRGTVKRSRYGQKL-----SMRTSGKL-----GARVHYQDLSPIGP----- 332
 QY 222 GDRKWGLASLTETLVCKELKLPNRIKGNWEPYPLSKOQLOVATDAYASWHLKYVKLD 281
 DB 333 -----TEDIEMEPILAP-----ASASAYDS--LYDYVADV 360

QY 282 PDVAVG 287
 DB 361 DDADIG 366

RESULT 10
 ID CPV1_MOUSE STANDARD; PRT; 503 AA.
 AC P28649;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cytochrome P450 19 (Aromatase) (EC 1.14.14.1) (CYPRX) (Estrogen synthetase) (P-450AROM).
 GN CYP19 OR AROM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=91378248; PubMed=1897929;
 RA Terashima M., Toda K., Kawamoto T., Kuribayashi I., Ogawa Y., Maeda T., Shizuta Y.;
 RT "Isolation of a full-length cDNA encoding mouse aromatase P450."
 RL Arch. Biochem. Biophys. 285:231-237(1991).
 CC - FUNCTION: CATALYZES THE FORMATION OF AROMATIC C18 ESTROGENS FROM C19 ANDROGENS.
 CC - CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
 CC - SUBCELLULAR LOCATION: Membrane-bound.
 CC - SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D00659; BAA00551.1; -
 DR PIR: S13912; S13912.
 DR MGD: MGI:88587; Cyp19.
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Electron transport; Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT BINDING 437
 SQ SEQUENCE 503 AA; 58015 MW; 737400389D1AEFF1 CRC64;

Query Match 6.1%; Score 91; DB 1; Length 503;
 Best Local Similarity 23.2%; Pred. No. 3.1;
 Matches 38; Conservative 23; Mismatches 33; Indels 70; Gaps 8;

QY 93 REGRLIYKSTATEDVNDKRAMQILKVLDTKRDESGIAFVGDIEM---RPSFRKGV-L-PGK 148
DB 115 REF-----SKRGLOICIGM-----HENGIIFFNNPNSLMRTIRPFMKALGPG 157
QY 149 VAVVOICVD-----SNVCYMHIFHSGIPDSLOHLED-STLVKVGIGD 192
DB 158 VRAVEVCESIKOHDLRGLGEVDTSGVDYLT-----MAHIMIDISNMLFLGIPLD 209
QY 193 GDSVK-----LFHDYGVSIKDVED 211
DB 210 ESAIVKIKGIQYFNAMQALLIKPNIFFKISWLYKRYERSVYDLMD 253
RESULT 11
DPOL_HAEIN STANDARD: PRT: 930 AA.
AC P43741;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL I).
GN POLA OR HI0856.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RA MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uiterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Georgagen N.S.M.,
RA Guelm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + (DNA)(N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U32767; AAC22515.1; -.
CC HSSP: P00582; IKFS.
CC TIGR: H10856; -.
DR InterPro: IPR002562; 3-5-exonuclease.
DR InterPro: IPR002421; 5-3-exonuclease.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR001098; DNA_pol_A.
DR InterPro: IPR000513; Exo_N.I.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF01612; 3-5-exonuclease; 1.
DR Pfam: PF01367; 5-3-exonuclease; 1.
DR Pfam: PF02739; 5-3-exonuclease; 1.
DR Pfam: PF00476; DNA_pol_A; 1.
DR PRINTS: PR00868; DNAPOL1.
DR SMART: SM00474; 35EXOC; 1.

DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; HHH2; 1.
DR SMART: SM00482; POLAC; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
SQ SEQUENCE 930 AA; 103740 MW; 22654BB7CEFF30B CRC64.
Query Match 6.0%; Score 89.5; DB 1; Length 930;
Best local Similarity 20.7%; Pred. No. 9.4;
Matches 61; Conservative 42; Mismatches 93; Indels 99; Gaps 15;
QY 24 ASYNFSRSSSSSAAPTVOATTSVHGHEDPNOIPNNIRQLPRSTTSSTKRPPLSR 83
DB 280 ARYEFRKWLNEVMNGADSIQT-----EOP-----YKNNQYKATSODQ 318
QY 84 CRARNFPAMRFGGRILYKSKATEVD-----KRAMQILVLTDRKDE-----SGIA 128
DB 319 SAVENTPKIQT-DRTKETILLQADLTWLEKLNAAKLAV-DTEIDSLDYSANLVGIS 376
QY 129 FVGLDIEWRPSFRKGVLPGRVATVOICVDSNYCDVMHIFHSGIPDSLOH-----LT 179
DB 377 F-----ALENGEAAYLPQLD--YLDA-----PKTLEKSTALAAIKPIL 413
QY 180 EDSTLVKVGIGIGDGYKLFHDYGVSIKDY-----DLSDLANOKIGCD 223
DB 414 ENPNHRIKIGNIKFE-SIPARHGIELQGVFPTMLSTYLNSTGRHMDLAKRYLG-- 470
QY 224 KKGGLASLPETLVCKELKPNRIRLGMWEFPLSKOOLQYAAADAVASMLYVL 278
DB 471 -----HETIAFESLAGKGSOL-TENOIPL-EQATEYAAEDADVTKMLQOAL 515
RESULT 12
MSH6_ARATH
ID MSH6_ARATH STANDARD: PRT: 1324 AA.
AC 004716;
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA mismatch repair protein MSH6-1 (AtMSH6-1).
GN MSH6-1 OR AGAA.3 OR AT4G02070 OR T10M13.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG_ERECTA;
RA TILL S., Granat S., Parnell L., Kaplan N., Hoffman J., Lodhi M.,
RA Johnson A.F., Dedhia N., Martienssen R., McCombe W.R.;
RL Subnucleotide (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=99444907; PubMed=10517319;
RA Ade J., Beizile F., Philippe H., Doutriaux M.P.;
RT "Four mismatch repair paralogues coexist in Arabidopsis thaliana:
RT AtMSH2, AtMSH3, AtMSH6-1 and AtMSH6-2.";
RL Mol. Gen. Genet. 262:239-249(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=20083488; PubMed=10617198;
RA Mayef K.F.X., Scheller C., Wandut R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terry N.,
RA Harris B., Anstorge W., Brandt P., Griwell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Maché R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Heineisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Billham L., Robben J.,

RA Van der Schueren J., Grymompres B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weizengger T., Bothé G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirse W.,
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA de Keyser A., Buysheert C., Gielen J., Villalroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
 RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Reichen S.,
 RA Borova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Faltmann B., Granderath K., Dauner D., Hertzl A.,
 RA Neumann S., Argitlou A., Vitale D., Liguori R., Pitravendi E.,
 RA Massena O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schabl S., Hiller R., Schmidt W., Lechary A., Aubourg S.,
 RA Chefor F., Cooke R., Berger C., Monfort A., Casachubeta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bleike C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dechta N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Frieden J.,
 RA Stoneking T., Kallick J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Speth J., Ryan E., Andrews S., Gelsel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shobdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana." 769-777(1999).
 RL Nature 402:769-777(1999).
 CC -1- FUNCTION: INVOLVED IN POST-REPLICATIVE DNA-MISMATCH REPAIR (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF001535; AAB57798.1; -
 DR EMBL: AJ245967; CAB53337.1; -
 DR EMBL: AF001308; AAC78699.1; -
 DR EMBL: AL161493; CAB80700.1; -
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR InterPro: IPR002999; Tudor.
 DR Pfam: PF00488; Muts_C.1.
 DR Pfam: PF01624; Muts_N.1.
 DR Prodom: PD001263; Muts_C.1.
 DR SMART: SM00534; Muts_C.1.
 DR SMART: SM00533; Muts_N.1.
 DR SMART: SM00333; Tudor.1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
 DR DNA repair: ATP-binding; DNA-binding.
 DR NP_BIND: 1083 1090 ATP (POTENTIAL).
 FT CONFLICT 316 318 ADR -> GPKSLLL (IN REF. 1).
 FT CONFLICT 383 383 O -> OVRRAHNL (IN REF. 1).
 FT CONFLICT 852 852 P -> PGITNLHLILILANTASHIISLP (IN
 FT REF. 1).
 SQ SEQUENCE 1324 AA; 146797 MW; 787A3A0272CE979C CRC64;

Query Match 6.0%; Score 89.5; DB 1; Length 1324;
 Best Local Similarity 23.2%; Predi No.15;
 Matches 68; Conservative 40; Mismatches 92; Indels 93; Gaps 18;
 QY 66 LPRSTSTSYKRP--LSRCRANPAMRGCG-----ILYSTATEVDRAMOLIKVL- 118
 DB 766 LPYSLEFRKSLRLPMEERLIARMFSSIEASGRNGKVLVEDPTA---RKQVDFISTLR 822
 QY 119 -----PRKR-----ESGIAVGDIDEMRPFRRG-VL 145
 DB 823 GCTTAAACSSLRALKHDTSLRLHLTPGOSLPNISSIKFKAFDVEVHNSGRVI 882
 QY 146 PKRVATVQICVDSNV---CDVMHIFHSIGIPQSH---LIEDSLVKVGIGID----- 192
 DB 883 PHEGA-----DEYDCACTVEEFESSLKHLKEGRKLIGDASINVTYVGNDEYLLLEV 936
 QY 193 -GDSYKLFHDY-----GVS-----IKDVELSLDANQIGDKKGLASTITELY-- 236
 DB 937 ESLSGVPDHYELCSSKSKSRVPTIKKL--LKLDSQAK--SEKESALKSISQRLGR 992
 QY 237 -CKELKPNRIRLGNMEFYPLSKOQLOVATDVAASMHLYKLDLPDVAVGS 288
 DB 993 FCHQEK-----WRLVSAIAEIDLVLISLAFAS-DTEGYRCRP-VISGS 1035
 RESULT 13
 BLM DROME STANDARD; PRT; 1487 AA.
 ID BLM DROME
 AC 09VG18: 09Y062;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Bloom's syndrome protein homolog (EC 3.6.1.-) (Dmbim) (Mutagen-
 sensitive protein 309) (RecQ helicase homolog).
 GN mus309 or blm or CG6920.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND REPEATS.
 RC STRAIN=Canton-S;
 RX MEDLINE=99160561; PubMed=10049920;
 RA Kusano K., Berres M.E., Engels W.R.;
 RT "Evolution of the RECQ family of helicases: a Drosophila homolog,
 RT Dmbim, is similar to the human Bloom syndrome gene.";
 RL Genetics 151:1027-1039(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brocktein P., Brothier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

FT CONFLICT      594      594      D -> G (IN REF. 1)
FT CONFLICT      605      614      G -> A (IN REF. 1)
FT CONFLICT      805      805      K -> E (IN REF. 1)
FT CONFLICT      1116     1116      E -> G (IN REF. 1)
SQ SEQUENCE      1487 AA; 166077 MW; 07361B8005E29432 CRC64;

Query Match          6.0%; Score 89.5; DB 1; Length 1487;
Best Local Similarity 24.5%; Pred. No. 18;
Matches    64; Conservative   42; Mismatches    94; Indels    61; Gaps    16;

QY      60 NNIR-ROLP-SITSTSYKRRPLSRRCRARNPPARFGRIYSTKATEVKKRAMQLIKV 117
       :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      930 SNLRRLVLPKKGVSFLTIDISRIKSK-----PMHFSG-IITYCSRKCEDETSMKKCK- 980
QY      118 LDRKDESGIAFV----GLDIEWRPSEFKGVLPGRV---ATVOICVDNSNYCDVMHIERS 169
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      981 -----DGVRASVSHAGLTPTDFRESQRKMVLTKMRIVATCAVGWGIDKPDVFVFLHY 1033
QY      170 GIPSLQHILIEDSTLYKKGIGIDGGSVR--LEHDYG--VSIKDVDDLSDLANQ-----KIG 221
       ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1034 SLPKSIEGYOEAGR-----GRDGDVADCIILYYNSDMIRIKMMD-SDKALQYNVKKI 1088
QY      222 GDKKGGASLTFETLV-CHELLKPNIIRLGNMFEFYLSKOOLQ-----YKAT 266
Db      1089 VDNLTRIYGCGNLTDCRRQA-----QLDYGEHPTSEQCLENRETACDNCINKRAYAV 1143
QY      267 DAYASMHLKYVLKDLPDAVSG 287
Db      1144 DALE--HARKAARAYKDLCSG 1162

RESULT 14
HK22__B_RARE
ID HK22_B_RARE STANDARD: PRT: 269 AA.
AC Q90J481:
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Homeobox protein NKX-2.2.
GN NKX2-2 OR NKX2.2 OR NK2.2.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
NCBI TaxID=7955;
RN [1]
RP MEDLINE FROM N.K.A.
RX MEDLINE=95324401; Pubmed=7600991;
RA Barth K.A., Wilson S.W.;
RT "Expression of zebrafish nkx2.2 is influenced by sonic hedgehog/vertebrate hedgehog-1 and demarcates a zone of neuronal differentiation in the embryonic forebrain."
RL Development 121:1755-1768(1995).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A CONTINUOUS NARROW BAND OF CELLS ALONG A BOUNDARY ZONE DEMARCATING THE LOCATION AT WHICH TWO OF THE EARLIEST NUCLEI IN THE BRAIN DIFFERENTIATE.
CC -!- SIMILARITY: BELONGS TO THE NK-2 FAMILY OF HOMEOBOX PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/or send an email to license@isb-sib.ch).
-----
CC EMBL; X85977; CAA59967.1; -.
CC HSSP; P22808; INK3.
DR ZFIN; ZDB-GENE-980526-403; nkx2.2.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox_1.
DR PRINTS; PR00024; HOMEBOX.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      October 31, 2002, 13:40:24 ; Search time 18 Seconds
              (without alignments)
              1537.429 Million cell updates/sec
```

Title: US-09-896-186B-24
Pagefoot: 1403

Sequence: 1 MSSSNWIDDAFTEELLAI...YASWHLKYVLKDLPAVSGS 288

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

database :
          pir_1: *
          1: pir1: *
          2: pir2: *
          3: pir3: *
          4: pir4: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1439.5	96.5	313	2	T05556	hypothetical prot
2	339.5	22.8	1436	2	T14895	hpa helicase 1 - 1
3	339	22.7	1401	2	T30247	werner syndrome p
4	330	22.1	1401	2	T17452	werner syndrome p
5	179.5	12.0	582	2	H96604	probable 3'-5' ex
6	158	10.6	123	2	E84752	hypothetical prot
7	136.5	9.2	123	2	S40930	hypothetical prot
8	127.5	8.6	239	2	A84777	hypothetical prot
9	112.5	7.5	217	2	T02548	hypothetical prot
10	101	6.8	876	2	T19246	hypothetical prot
11	97.5	6.5	445	2	E72674	hypothetical prot
12	96.5	6.5	488	2	F97039	hypothetical prot
13	95	6.4	508	1	A36121	aromatase (EC 1.1
14	95	6.4	4589	1	T14914	aromatase (EC 1.1
15	94.5	6.3	416	2	C71620	dynein beta heavy
16	93.5	6.3	1525	1	GNWVS	protein with Egl-1
17	93	6.2	2224	1	KEHVS	genome polyprotein
18	93	6.2	3079	1	RGBY12	coagulation factor
19	92	6.2	237	2	H84732	probable GRPase-a
20	92	6.2	758	2	T39210	hypothetical prot
21	91.5	6.1	473	2	S36553	3-isopropylmalate
22	91.5	6.1	781	2	T41551	l2 protein human
23	91	6.1	503	2	S13912	hypothetical prot
24	90	6.0	472	2	T47436	aromatase (EC 1.1
25	89.5	6.0	930	2	E64098	protein kinase-11
26	89.5	6.0	1324	2	T01508	DNA-directed DNA
27	88	5.9	269	2	I50504	mismatch repair e
28	88	5.9	317	2	A97511	nk2.2 protein - 2
29	88	5.9	317	2	AH2729	agmatinase [impor

30	88	5.9	419	2	E82095	gamma-glutamyl phosphate reductase
31	88	5.9	707	2	S40926	hypothetical protein
32	88	5.9	784	2	C88558	protein ZK1098.3
33	87.5	5.9	332	2	T03556	probable D-ribose-5-phosphate 4-epimerase
34	87.5	5.9	422	2	E96753	hypothetical protein
35	87.5	5.9	547	2	T40342	signal recognition particle
36	87.5	5.9	589	2	E88492	protein T0E3.1
37	87	5.8	178	2	E72450	hypothetical protein
38	87	5.8	392	2	T43362	acetyl-CoA C-acetyltransferase
39	87	5.8	579	2	I40371	collagen alpha 3(V)
40	87	5.8	3337	1	A37797	phosphate-repressible protein
41	86.5	5.8	346	1	BYECPR	hypothetical protein
42	86.5	5.8	360	2	T00882	glycerol-3-phosphate dehydrogenase
43	86.5	5.8	470	2	T10193	hypothetical protein
44	86.5	5.8	496	2	S61966	hypothetical protein
45	86.5	5.8	593	2	B85844	hypothetical protein

ALIGNMENTS

RESULT 1
T05256
Hypothetical protein F18A5_260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05256
R:Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; M
Submitted to the Protein Sequence Database, February 1999
A:Reference number: 215405
A:Accession: T05256
A:Molecule type: DNA
A:Residues: 1-313 <BEV>
A:Cross-references: EMBL:AL035528
A:Experimental source: Cultivar Columbia; BAC clone F18A5-
C:Genetics:
A:Map position: 4
A:Introns: 86/1; 143/1; 186/3; 239/3
A:Note: F18A5_260

Query Match	Similarity	Score	DB 2:	Length
Best Local	95.3%	1439.5	DB 2:	313
Matches	201	Conservative	1	Mismatches 4; Indels 9; Gaps 1

Query	1	MSSSNMIDDAFT	EEELALDAI	EAASYNFSRSSSSSSAAPTQAATTSVHGHEEDPN	IPN	60
Db	1	MSSSNMIDDAFT	EEELALDAI	EAASYNFSRSSSSSSAAPTQAATTSVHGHEEDPN	IPN	60
Qy	61	NIRQLPRSTSTSTSKRRPLSRCAARNP	PMRFGRLYKTKAT	EVDRKAMOLIKYLD	120	
Db	61	NIRQLPRSTSTSTSKRRPLSRCAARNP	PMRFGRLYKTKAT	EVDRKAMOLIKYLD	120	
Qy	121	KRDESGIAEFGDLIEWRPSFRKGVLP	RGVATQLCVDSNCDVMNHGSI	POSTQHLIE	180	
Db	121	KRDESGIAEFGDLIEWRPSFRKGVLP	RGVATQLCVDSNCDVMNHGSI	POSTQHLIE	180	
Qy	181	DSTLVAVGIGIDGDSYKLFHDYGV	STIKVEDISDLANOK	IGDGKWKGLASTETPLVCKEL	240	
Db	181	DSTLVAVGIGIDGDSYKLFHDYGV	STIKVEDISDLANOK	IGDGKWKGLASTETPLVCKEL	240	
Qy	241	LKPNRIRLGNMEFPYLSKQOLQY	AAATDVAASWHLTKV-----	LKDLPDAYS	286	
Db	241	LKPNRIRLGNMEFPYLSKQOLQY	AAATDVAASWHLTKV-----	LKDLPDAYS	286	

RESULT	2
T14895	DNA helicase 1 - African clawed frog
C:Species:	Xenopus laevis (African clawed frog)
C:date:	20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession:	T14895
R:Yan, H.; Chen, C.Y.; Kobayashi, R.; Newport, J.	Nature Genet. 19, 375-378, 1998

A:Title: Replication focus-forming activity 1 and the Werner syndrome gene product.

A:Reference number: Z18255; MUID:98361165

A:Accession: T14895

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1436 <YAN>

A:Cross-references: EMBL:AF067418; NID:g3420290; PID:g3420291; PIDN:AAC63512.1

C:Genetics:

A:Gene: FFA-1

Query Match 22.8%; Score 339.5; DB 2; Length 1436;

Best Local Similarity 35.0%; Pred. No. 8.6e-21;

Matches 82; Conservative 43; Mismatches 84; Indels 25; Gaps 7;

Db 3 SLORLPEMWSVQOEDRIDDAKSKFKNILEDNLPKFKNSIYYSNCCSLSD 62

Qy 61 NIPROLPRSTSTSYKRF---PLSRCRAR---NFPAMFGRIILYSKATEV-----D 108

Db 63 IRSSLL-----EDV---LGFDIEMPPVYTKG-KTGKVALIOVCSEKCYLFPHIS 110

Qy 109 KRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGVLPCKVATVOICVDSNYCDVMHIF- 167

Db 166 HSGIPSOHLIEDSTLVYVGIGIDGDSVKLFPHDYGVSTIKDVDDLSDLANOKIGGDKWG 227

Qy 111 MAFEPGKLRLLEDSEVRKVGVEGDQWKLMSDYELKLGTELEMANOKLRCKEKT 170

Db 226 LMSLTFLVCKELLPNIRIRLGNMEFVPLSKOOLQYAAATDAYASWHLVKVLDL 281

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Qy 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

Db 171 FNLILHLEFREQYLRKRSYCSNMDIFLTEDOKLVAATDAYAGLLYKRLDEGM 224

RESULT 4

T17452

Werner syndrome protein - mouse

N:Alternate names: wrn protein

C:Species: Mus musculus (house mouse)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17452

R:Paepker, B.W.; Gayle, M.; Brady, W.; Swartz, A.; Gillett, L.A.; Allisch, R.S.; Mullig

submitted to the EMBL Data Library, September 1998

A:Description: Genomic structure of the human Werner's gene and cloning of its mouse

A:Reference number: Z18794

A:Accession: T17452

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1401 <PAR>

A:Cross-references: EMBL:AF091215; NID:g3885837; PID:g3885838; PIDN:AAC78077.1

C:Genetics:

A:Gene: wrn

Query Match 22.1%; Score 330; DB 2; Length 1401;

Best Local Similarity 36.0%; Pred. No. 5.5e-20;

Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;

Db 88 NFPAMFGRIILYSKATEYDKRAMOLIKVLDTRKDESGIAFVGLDIEMRPSFRKGVLP 147

Qy 37 NLPFLFPGSIIVSYEASD---CSFLSEDISMRSLSDGV--VGFMEMPPIYK----Pg 86

Db 148 K--VATVOICVDSNYCDVMHIFHSGI-POSTLOHLEDSTLVYVGIGIDGDSVKLFPHDY 203

Qy 87 KRSRAVAVIQLCVSESKCYLFPHISMSVFPQGLKMLENKSIRKAGVIGEDQKLRD 146

Db 204 VSIKDVEDSLANOKIGGDKKGLASLTETLVCKELLPNIRIRLGNMEFVPLSKOOLQ 263

Qy 147 VKLESFVELTDVANERKLCALFTWMSLVGLVKQLKDKSRCSNWSNFPITEDQKLY 206

Db 264 AATDAYASWHLVKVLDLP 283

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Qy 207 AATDAYAGLLIYKRLGND 226

Db 207 AATDAYAGLLIYKRLGND 226

Db 43 DLPELEFTGSIYVSYDASCSFLSE-----DISMSLSDDGVVCFDMEWPPLYNRGL-G 95
Qy 148 KVAATVOICVSNCDVWHIFHSGI-POSTLOHIEDSTLVVVGIGIDGDSVKLFHDYCSI 206
Db 96 KVALIQCVSESKCYLTHVSMYSVFPQGLMLLENKRAVGAIGIEDQKLLRDFIKL 155
Qy 207 KDVEDLSDLANOKIGDKKGLASLTFTLVCKELLPNRIRLGNWFEYPLSKOOLQYAA 266
Db 156 KNFVELDVANKKCLKCTETMSLNSLVKHLGKQLLKKSIRCSNWSKFPLEEDQKLYLAT 215
Qy 267 DATDAYASHMLYKVLKDLDPAY 285
Db 216 DAYAGFTIYRNLELIDDTV 234

RESULT 2
US-09-127-670-6
Sequence 6, Application US/09127670
Patent No. 6228583
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute for Technology
APPLICANT: Leonard P. Guarente
APPLICANT: David A. Sinclair
APPLICANT: David B. Lombard
TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE
TITLE OF INVENTION: SPAN
FILE REFERENCE: MIT-7720PA
CURRENT APPLICATION NUMBER: US/09/127,670
CURRENT FILING DATE: 1998-07-31
EARLIER APPLICATION NUMBER: US 60/054,629
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1401
TYPE: PRM
ORGANISM: Murine
US-09-127-670-6

Query Match 22.7%; Score 339; DB 4; Length 1401;
Best Local Similarity 36.6%; Pred. No. 2e-28;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;

Qy 88 NFPMAREGRILYSKATVEYDKRAMOLIKVLDTRKDESGIAVGLDIEMRPSFRKGYLP 147
Db 37 NLPELEFGSIYVSYSASD-----CSFLSEDISMRSLSDGV--VGDFMEWPPIYK----PG 86
Qy 148 K---VAATVOICVSNCDVWHIFHSGI-POSTLOHIEDSTLVVVGIGIDGDSVKLFHDY 203
Db 87 KRSRAVAITQCVSESKCYLTHVSMYSVFPQGLMLLENKSIKRAVGAIGIEDQKLLRDF 146
Qy 204 VSIKDVEDLSDLANOKIGDKKGLASLTFTLVCKELLPNRIRLGNWFEYPLSKOOLQY 263
Db 147 VKLESFVELDVANENKKAETWSLNGLVKHLGKQLLKKSIRCSNWSKFPLEEDQKLY 206
Qy 264 AATDAYASHMLYKVLKDLDPAY 285
Db 207 AATDAYAGLITYOKLGNLGDV 228

RESULT 3
US-08-781-891-206
Sequence 206, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209

Correspondence Address:
Addressee: Seed and Berry LLP
Street: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
State: Washington
Country: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 206:
SEQUENCE CHARACTERISTICS:
LENGTH: 1401 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-891-206

Query Match 22.1%; Score 330; DB 3; Length 1401;
Best Local Similarity 36.0%; Pred. No. 2e-27;
Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;

Qy 88 NFPMAREGRILYSKATVEYDKRAMOLIKVLDTRKDESGIAVGLDIEMRPSFRKGYLP 147
Db 37 NLPELEFGSIYVSYSASD-----CSFLSEDISMRSLSDGV--VGDFMEWPPIYK----PG 86
Qy 148 K---VAATVOICVSNCDVWHIFHSGI-POSTLOHIEDSTLVVVGIGIDGDSVKLFHDY 203
Db 87 KRSRAVAITQCVSESKCYLTHVSMYSVFPQGLMLLENKSIKRAVGAIGIEDQKLLRDF 146
Qy 204 VSIKDVEDLSDLANOKIGDKKGLASLTFTLVCKELLPNRIRLGNWFEYPLSKOOLQY 263
Db 147 VKLESFVELDVANENKKAETWSLNGLVKHLGKQLLKKSIRCSNWSKFPLEEDQKLY 206
Qy 264 AATDAYASHMLYKVLKDLDPAY 283
Db 207 AATDAYAGLITYOKLGNLGDV 226

RESULT 4
PCT-US94-00198-4
Sequence 4, Application PC/TUS9400198
GENERAL INFORMATION:
APPLICANT: Schering Corp.
TITLE OF INVENTION: RAS Associated GAP Proteins
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 1 Giraldo Farms
City: Madison
State: New Jersey
Country: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 6.0.8
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00198

```

: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/004,824
: FILING DATE: 15-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Lunn, Paul G.
: REGISTRATION NUMBER: 32,743
: REFERENCE/DOCKET NUMBER: DX0352 PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (201)822-7255
: TELEFAX: (201)822-7039
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3079 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Saccharomyces cerevisiae
: PCT-US94-00198-4

```

Query Match 6.2%; Score 93; DB 5; Length 3079;

Best Local Similarity 19.3%; Pred. No. 1.7; Indels 84; Gaps 10;

Matches 57; Conservative 43; Mismatches 112; Indels 84; Gaps 10;

```

QY 9 DAFTFEEELAIIDAIEASYNFSRSSSSAAPTVOATTSVHGHEEDPNQIPNNIRQLPR 68
DB 1976 DDFYKTFLLIDVYLGOLQPKMEFNS-----ETPIYREMDYDPELYEFMNNHAFR 2028
QY 69 SITSSSY-----KRPPLSRCRRNPMPMRFGRLYSKTATENDKRAMOLIK-- 116
DB 2029 NIETSTAYSPSVHESTSESEIPITLTMSNF-----SDRHIDIDTVAYKFLQIY 2077
QY 117 -----VLDTKRDESGIAFVGLDIEMRPSFRKGVLPGRKATVQICDSVYCDVMHI 166
DB 2078 ARWTKRHCLIDCTERDEG---GLDMKRFEISLVWGGLP-EVAP-KNCIGCYFPVNET 2131
QY 167 F-----HSGIPQSLQHLIEDSTLVKVGIGIDGDSVKLFHDYGSIRKVED 211
DB 2132 FMNNGKCLDKDNVYSSKPHYFINSNDEGLMK-SVGTGGGLAVLQDIRSLMDITL 2190
QY 212 LSLNALOKIGGKKKGLASITFLVCKELKPNRIRLGNMEFYPLSKOOLQVATD 267
DB 2191 YDEKRRN-----FTPVSLKIGDIYFOVLHETPROXYKIRD 2224

```

RESULT 5

US-09-722-139-2
Sequence 2, Application US/09722139
Patent No. 6355471

GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6355471el motor proteins and methods for
FILE REFERENCE: 1055
CURRENT APPLICATION NUMBER: US/09/722,139
CURRENT FILING DATE: 2000-11-24
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1375
TYPE: PRT
ORGANISM: Human
US-09-722-139-2

Query Match 5.9%; Score 87.5; DB 4; Length 1375;

Best Local Similarity 19.0%; Pred. No. 2;

Matches 58; Conservative 52; Mismatches 116; Indels 79; Gaps 13;

QY 13 EEELAIIDAIEASYNFSRSSSSAAPTVOATTSVHGHEEDPNQIPNNIRQLPRSTS 72

```

DB 935 KEEOLA-----QYQANANOLOKLOATTEFTANIAROE-----KVRKKEKILES 979
QY 73 STYKRPPLSRCRRNPMPMRFGRLYSKTATEVDKRAMOLIKVLDTKRDESGIAFVGL 132
DB 980 REKOQREALERALAR--LEKRHSALQRHSLTGLTEIFEQOKLASLMSGSRREGSLQ-ASL 1036
QY 133 DIEMRPSFRK-----GVLPGKRVATVQICVDSNYCDVMH 165
DB 1037 EAE-QEALKEQDERLEYEQLOKQIYEVGVQKDHGTLLEGKVASSSLPVSAEKSHLVP 1095
QY 166 I-----FHSGIPQSLQHLIED-STLVKVGIGIDGDSVKLFHDYGVSTIKVDLSDLANOKI 220
DB 1096 LMDARINAYIEEVOQRLODLHRIYISEGCSSTADTK-----DNEKLNHGTI 1142
QY 221 GGDKKWGLASITFLVCKELKPNRIRLGNMEFYPLSKOOLQVATDAYASWHLVYKLD 280
DB 1143 QRLKLY---ELCRDLICVLMPEPDAACAN---HPLLQDLVQSLD---WK-----TE 1187
QY 281 LPDAV 285
DB 1188 IPLDV 1192

```

RESULT 6

US-07-989-845-2
Sequence 2, Application US/07989845
Patent No. 5304472

GENERAL INFORMATION:
APPLICANT: Bass, Steven
APPLICANT: Swartz, James
TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE
NUMBER OF INVENTION: PRODUCTION IN BACTERIAL CELLS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080-4990
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/989,845
FILING DATE: 19921120
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 752
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-989-845-2

Query Match 5.8%; Score 86.5; DB 1; Length 348;

Best Local Similarity 20.9%; Pred. No. 0.3;

Matches 54; Conservative 34; Mismatches 73; Indels 97; Gaps 12;

QY 86 ARNFPAMRFGRLYSKTATEV-----DKRAM-----OLIKVLDTKRDESG 126

```

Db 107 AVNIPGLKSGELVDGKLTGDIYLGKIKKWDDEAIKLNPGKLPSONIAV--RRADGSG 165
QY 127 IAFV-----GDIEMRPSRKGV-----LPGVNAVQICVDS 158
Db 166 TSVFTSYLAKVVEKNNVGTSTVKWPIGLGKNGNDIAIAVORLPGAIGVEYA-- 222
QY 159 NYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKVEDLS----- 213
Db 223 -YAKONNLAVTKL-----ISADGKPVSPSTEENFANNAKAGADMSTFQ 264
QY 214 DLANOKIGDDKWKGLASLETETVCKELKPNR-----IRLGNWEFPLSKO--OLOYAATD 267
Db 265 DLTNOK--GEDAMPITSTFILLHKDOKKPEQGEVLKFFDMAVYKTGAKOANDLDYAS-- 320
QY 268 AYASWHLVYKVLDPDAV 285
Db 321 -----LPDSV 325

```

RESULT 7

PCT-US93-11298-2

```

; Sequence 2, Application PC/TUS9311298
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE PRODUCTION IN
; TITLE OF INVENTION: BACTERIA
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: palin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11298
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 752
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US93-11298-2

```

Query Match 5.8%; Score 86.5; DB 5; Length 348;
 Best Local Similarity 20.9%; Pred. No. 0.3;
 Matches 54; Conservative 34; Mismatches 73; Indels 97; Gaps 12;

```

QY 86 ARNFPMRFGRLVSKTATEV-----DKRAM-----OLIKVLDTKRDESG 126
Db 107 AVNIPGLKSGELVDGKLTGDIYLGKIKKWDDEAIKLNPGKLPSONIAV--RRADGSG 165
QY 127 IAFV-----GDIEMRPSRKGV-----LPGVNAVQICVDS 158
Db 166 TSVFTSYLAKVVEKNNVGTSTVKWPIGLGKNGNDIAIAVORLPGAIGVEYA-- 222

```

```

QY 159 NYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGSVKLFHDYGVSIKVEDLS----- 213
Db 223 -YAKONNLAVTKL-----ISADGKPVSPSTEENFANNAKAGADMSTFQ 264
QY 214 DLANOKIGDDKWKGLASLETETVCKELKPNR-----IRLGNWEFPLSKO--OLOYAATD 267
Db 265 DLTNOK--GEDAMPITSTFILLHKDOKKPEQGEVLKFFDMAVYKTGAKOANDLDYAS-- 320
QY 268 AYASWHLVYKVLDPDAV 285
Db 321 -----LPDSV 325

```

RESULT 8

US-08-907-166-12

```

; Sequence 12, Application US/08907166
; Patent No. 5948666
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/08/907,166
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO: 12
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Aquifex pyrophillus
; US-08-907-166-12

```

Query Match 5.4%; Score 80; DB 2; Length 574;
 Best Local Similarity 21.9%; Pred. No. 3.5;
 Matches 33; Conservative 30; Mismatches 48; Indels 40; Gaps 6;

```

QY 170 GIPQSLQHLIEDSTLVKVGIGIDGSVKLF-----HDYGV--SIKVEDLSLANOK- 219
Db 11 GLKKAIRL-ENSPYLYDTEETGDRIRLVQIGDEENTVIDLYEIDIEPLKILNERG 69
QY 220 -IGDKKKGL-----ASLETETLVCKELKPNRRL-----GN 250
Db 70 YGHNILKFDLKYLRGIPPSATFFDMIASYLLGYERHSLNIHVSNLGYSMDKSYOTSD 129
QY 251 WEFPYLSKQLOLYAATDAVASWHLVYKVLKDL 281
Db 130 WGNASVLSDAQLKTAANDVYLVRELFPKMRDM 160

```

RESULT 9

US-08-845-258-34

```

; Sequence 34, Application US/08845258
; Patent No. 6183976
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

```

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS


```

; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-723-142A-34

```

Query Match	5.3%	Score 79.5	DB 4	Length 431
Best Local Similarity	22.4%	Pred. No. 2.5		
Matches 66	Conservative 35	Mismatches 106	Indels 87	Gaps 17

[illegible]

```

RESULT 12
US-08-895-601-6
: Sequence 6, Application US/08895601
: Patent No. 6060262
:
: GENERAL INFORMATION:
: APPLICANT: Beer-Romero, Peggy
: APPLICANT: Strack, Peter J.
: APPLICANT: Glass, Susan J.
: APPLICANT: Rolfe, Mark
: TITLE OF INVENTION: REGULATION OF KAPPA B (1KB) DEGRADATION,
: TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
: NUMBER OF SEQUENCES: 16
:
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: FOLEY, HOAG & ELIOT LLP
: STREET: One Post Office Square
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109-2170
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/895,601
: FILING DATE: 16-JUL-1997
: CLASSIFICATION: 435
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: MIT-096.01
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-832-7000
: TELEFAX: 617-832-7000
:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 927 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-08-895-601-6

```

Query Match	5.3%;	Score 79;	DB 3;	Length 927;
Best Local Similarity	20.9%;	Pred: No. 9.5;		
Matches	63;	Conservative	37;	Mismatches 82;
				Indels 120;
				Gaps 19

OY	22	LEASVNFSSSSSSSAPVQATSVNGH-EDPNOPNNIRROLPRSTISSTYKRP	80
Dp	408	VQATVEITSQLISSOSSANQPOQSATSDBQQVOVPSEIEGCF--LRKG-----	455
OY	81	LSRCARFPRMRRGRGLLY----SKRATEVKRAMOLIKY-----LDTKRD----	128
Dp	454	--MEVRHAIP--NGREFFIDHNKTFTTWDDPR---LKRAHLRKGTSLDTSNDIGPL	502
OY	124	-----ESGIAFVOLDI---EWR-PSEFKGVLPDKVATVOICVDNSCYDMHIIF-	167
Dp	503	PWGMEERTHTDGRIIFYUINHNIKQTOWMEDPRLFNALNGRA---VPYSRDYKKRYEFFRR	558
OY	168	-----HSGRPSQLQ-----HLIEDS-----TLVKGIGIGDSVLFHDYGVSIK	207
Dp	559	KLKRONDIRPKFEKKLRATVALEDSTSYRRIKMKRADFLKARTLWTFEGEGL-DYG---	613
OY	208	DVEDLDLANOKIGGDKKWGLASTLETIVCKELKPNNIRILGNMFPLSKOOLQYAADT	267
Dp	614	-----GVAREWMVF-----LISKEMENP-----YYGL-----PEYSATD	644
OY	268	AY 269	
Dp	642	NY 643	

RESULT 13
 US-07-640-029-2
 Sequence 2, Application US/07640029
 Patent No. 5229501
 GENERAL INFORMATION:
 APPLICANT: Kiefer, Michael C.
 APPLICANT: Valenzuela, Pablo D.T.
 APPLICANT: Barr, Philip J.
 TITLE OF INVENTION: Expression and Use of Human Fibroblast
 TITLE OF INVENTION: Growth Factor Receptor
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chilton Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: California
 COUNTRY: USA
 ZIP: 94608
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-07/640,029
 FILING DATE: 19910111
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: CH-165
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 510-601-2708
 TELEFAX: 510-655-3542
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 817 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: Single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-07-640-029-2

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 31, 2002, 13:36:19 ; Search time 34 Seconds
(without alignments)
940.860 Million cell updates/sec

Title: US-09-896-186b-24

Perfect score: 1491

Sequence: 1 MSSSMWIDAFTEELLAIID.....YASWHLKVKLDLPDAVSGS 288

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1491	100.0	288	21	AAAG52858
2	1491	100.0	298	21	AAAG52857
3	1491	100.0	298	21	AAAG52856
4	1028	68.9	197	21	AAAG52855
5	986	66.8	197	21	AAAG52854
6	896	60.1	177	21	AAAG52853
7	352	23.6	1432	18	AAAT14517
8	339	22.7	1401	19	AAAW59454
9	339	22.1	1401	20	AAAW59453
10	330	22.1	1401	18	AAAT14519
11	229.5	15.4	346	22	AAAB71575

12	212.5	14.3	583	22	AAAB62653	Drosophila melanog
13	178.5	12.0	244	22	AAAB80638	Environmental stre
14	151	10.1	127	22	AAAB11264	Human secreted pro
15	151	10.1	147	20	AAAY93336	Human secreted pro
16	151	10.1	147	22	AAAY93087	Human secreted pro
17	151	10.1	466	22	AAAB93116	Human protein sequ
18	105	7.0	625	22	AAAB58154	Drosophila melanog
19	103	6.9	220	21	AAAG04684	Arabidopsis thalia
20	96	6.4	706	16	AAAR75647	Thermophilic bacte
21	95	6.4	161	21	AAAG28363	Arabidopsis thalia
22	95	6.4	167	21	AAAG28362	Arabidopsis thalia
23	94.5	6.3	416	21	AAAB18185	Novel human diagno
24	93	6.2	2224	17	AAAG04254	Human Factor V. H
25	93	6.2	2224	20	AAAY49564	Human lipoprotein
26	93	6.2	451	21	AAAG18364	Arabidopsis thalia
27	91	6.1	477	21	AAAB18363	Arabidopsis thalia
28	91	6.1	477	21	AAAB18363	Arabidopsis thalia
29	89.5	6.0	346	15	AAAB60653	psts variant. Esc
30	89.5	6.0	930	22	AAAG35540	Haemophilus influe
31	89.5	6.0	1078	22	AAAB61252	Drosophila melanog
32	89.5	6.0	1487	22	AAAB62792	Drosophila melanog
33	89	6.0	991	21	AAAY83171	Cell wall protein
34	89	6.0	991	21	AAAY70120	Staph. epidermidis
35	88.5	5.9	3079	15	AAAB59926	GAP protein Iraz.
36	88	5.9	506	22	AAAB64451	Drosophila melanog
37	87.5	5.9	642	22	AAAB61316	Drosophila melanog
38	87.5	5.9	2354	22	AAAB60511	psts variant. Esc
39	86.5	5.8	346	15	AAAB60640	psts variant. Esc
40	86.5	5.8	346	15	AAAB60641	psts variant. Esc
41	86.5	5.8	346	15	AAAB60642	psts variant. Esc
42	86.5	5.8	346	15	AAAB60643	psts variant. Esc
43	86.5	5.8	346	15	AAAB60644	psts variant. Esc
44	86.5	5.8	346	15	AAAB60645	psts variant. Esc
45	86.5	5.8	346	15	AAAB60646	psts variant. Esc

ALIGNMENTS

RESULT 1

ID AAAG52858 standard; Protein; 288 AA.

XX AC AAAG52858;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67237.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 9905-0121825.

XX PR 05-MAR-1999; 9905-0123180.

XX PR 09-MAR-1999; 9905-0123548.

XX PR 23-MAR-1999; 9905-0125788.

XX PR 25-MAR-1999; 9905-0126264.

XX PR 29-MAR-1999; 9905-0126785.

XX PR 01-APR-1999; 9905-0127462.

XX PR 06-APR-1999; 9905-0128234.

XX PR 08-APR-1999; 9905-0128714.

XX PR 16-APR-1999; 9905-0129845.

XX PR 19-APR-1999; 9905-0130077.

XX PR 21-APR-1999; 9905-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143452.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1491; DB 21; Length 288;
Best Local Similarity 100.0%; Pred. No. 4,6e-151;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNMWIDAFTEEEELAIIDAIEASYNFSRSSSSSAAPTVOATTSVHGHEEDPNQIPN 60
DB 1 MSSNMWIDAFTEEEELAIIDAIEASYNFSRSSSSSAAPTVOATTSVHGHEEDPNQIPN 60
QY 61 NTRQLPRSTSTSTSKRRPLSKRCRANRPNMFGGRILYSKATIEVDKRAMOLIVLDT 120
DB 61 NTRQLPRSTSTSTSKRRPLSKRCRANRPNMFGGRILYSKATIEVDKRAMOLIVLDT 120
QY 121 KRDESGIATVGLDIEWRPFGRKGVLPKATVQICVDSNCDVMHIFHSGIPQSLQHLIE 180
DB 121 KRDESGIATVGLDIEWRPFGRKGVLPKATVQICVDSNCDVMHIFHSGIPQSLQHLIE 180
QY 181 DSTLVKVGIGIDGDSVKLFDHYGVSTIKDVEDLSDLANOKRIGGDKKMGSLATETLVCKEL 240
DB 181 DSTLVKVGIGIDGDSVKLFDHYGVSTIKDVEDLSDLANOKRIGGDKKMGSLATETLVCKEL 240
QY 241 LKPNRIRLGNWFEYPLSKQOLQYATDAYASWHLIKVLDLPDAVSGS 288
DB 241 LKPNRIRLGNWFEYPLSKQOLQYATDAYASWHLIKVLDLPDAVSGS 288

RESULT 2

AAG52857 standard; Protein; 298 AA.

AAG52857;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 67236.

Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EPI033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155859.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1491; DB 21; Length 298;
Best Local Similarity 100.0%; Pred. No. 4, 9e-151;

Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSNWIDDAFTEEEELAIIDAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNOIPN 60
DB 11 MSSSSNWIDDAFTEEEELAIIDAIEASYNFSSSSSSSAFTVOATTSVHGHEEDPNOIPN 70
QY 61 NRRQLPRSTTSSTSYKRFPLSRCRANRPNRPFGRILYKATVEYDKRAMOLIVLDT 120
DB 71 NRRQLPRSTTSSTSYKRFPLSRCRANRPNRPFGRILYKATVEYDKRAMOLIVLDT 130
QY 121 KRDESGIAFYGLDIEMRPFRKGVLPCKVATVOICVDNVCYVHMLFHSGLPQSLQHLIE 180
DB 131 KRDESGIAFYGLDIEMRPFRKGVLPCKVATVOICVDNVCYVHMLFHSGLPQSLQHLIE 190
QY 181 DSTLVKVGIGIDGDSVKLFHDYGVSTIKDVEDLSDLANOKTGGKKNGLASTTETLVCKEL 240
DB 191 DSTLVKVGIGIDGDSVKLFHDYGVSTIKDVEDLSDLANOKTGGKKNGLASTTETLVCKEL 250
QY 241 LKPNRIRLGNMEFYPLSKOOLYAAATDAYASWMLYKVLKDLPPAVSGS 288
DB 251 LKPNRIRLGNMEFYPLSKOOLYAAATDAYASWMLYKVLKDLPPAVSGS 298

RESULT 3

AAAG24977

ID AAAG24977 standard; Protein; 288 AA.

XX AAAG24977;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28856.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-01231825.
PR 05-MAR-1999; 99US-01223180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 96.1%; Score 1433; DB 21; Length 288;
Best Local Similarity 96.5%; Pred. No. 7.6e-145;

Matches 278; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSSSNWIDAFTEEEELALDAIEASYNFSRSSSSSAFTVOATTSVHGHEEDPNQIPN 60
DB 1 MSSSNWIDAFTEEEELALDAIEASYNFSRSSSSSAFTVOATTSVHGHEEDPNQIPN 60
QY 61 NNRQLPRSTTSSTYKRFFLSRCRANFPAMRFGGRILYSKTATEVDKRAMOLIKVLD 120
DB 61 NNRQLPRSTTSSTYKRFFLSRCRANFPAMRFGGRILYSKTATEVDKRAMOLIKVLD 120
QY 121 KRDESGIAFYGLDIEMRPSRKGVLPCKVATVQICVDNSNCDVMHIFHSGIPQSLQHLIE 180
DB 121 KRDESGIAFYGLDIEMRPSRKGVLPCKVATVQICVDNSNCDVMHIFHSGIPQSLQHLIE 180
QY 181 DSTLVKVGIGIDGSYKLFPHDYGVSIKDYEDLSDLANOKIGDKKMGCLASTETLVCKEL 240
DB 181 DSTLVKVGIGIDGSYKLFPHDYGVSIKDYEDLSDLANOKIGDKKMGCLASTETLVCKEL 240
QY 241 LKPNRIRLGNMEFYPLSKOOLQYAATDAYASWHLKYVKLDPDAVSDS 288
DB 241 LKPNRIRLGNMEFYPLSKOOLQYAATDAYASWHLKYVKLDPDAVSDS 288

RESULT 4

AA652859

ID AA652859 standard; Protein: 197 AA.

AC AA652859;

XX 18-OCT-2000 (first entry)

BT 18-OCT-2000 (first entry)

XX XX Arabidopsis thaliana protein fragment SEQ ID NO: 67238.
DE DE
XX XX Protein identification: signal transduction pathway; metabolic pathway;
KW KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW KW termination sequence.
XX XX Arabidopsis thaliana.
OS OS
PN PN EP1033405-A2.
PD PD
XX XX 06-SEP-2000.
PF PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130077.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132047.
PR 30-APR-1999; 9905-0132408.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 17-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.

```
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140921.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.9%; Score 1028; DB 21; Length 197;
Best Local Similarity 100.0%; Pred. No. 1,le-101;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 MRGGRIILYSKTATEVDKRAMOLIKVLDTRKRDSSGIAFPGLDIEKMPRSRKVLPKVAFT 151
DB 1 MRGGRIILYSKTATEVDKRAMOLIKVLDTRKRDSSGIAFPGLDIEKMPRSRKVLPKVAFT 60
QY 152 VOICVSNYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGDSVYLFHDYGVSIKDVED 211
DB 61 VOICVSNYCDVWHIFHSGIPQSLQHLIEDSTLVKVGIGIDGDSVYLFHDYGVSIKDVED 120
QY 212 LSDLANOKTIGDKKMWGLASITETLVCKELLKPNRIRLGNWFEYPLSKOOLQYAAATDAVAS 271
DB 121 LSDLANOKTIGDKKMWGLASITETLVCKELLKPNRIRLGNWFEYPLSKOOLQYAAATDAVAS 180
QY 272 WHLYKVLKDLPPAVSGS 288
DB 181 WHLYKVLKDLPPAVSGS 197

RESULT 5
AAG24978
```

ID AAC24978 standard; Protein; 197 AA.
XX
XX AAG24978;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28857.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI03405-A2.
XX
XX 06-SEP-2000.
PD
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.8%; Score 996; DB 21; Length 197;
Best Local Similarity 97.0%; Pred. No. 2.9e-98;
Matches 191; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 92 MRGGRILYKTEVDEKRAMOLIKVLDTRKDSGIAFVGLDEMRPSFKVLPGKVFAT 151
DB 1 MRGGRILYKTEVDEKRAMOLIKVLDTRKDSGIAFVGLDEMRPSFKVLPGKVFAT 60
OY 152 VOICVDNNTCDVMHIFHSGIPQSLQHLIEDSTLVKVGIGIDGDSVKLFDHYGVSIXKVED 211
DB 61 VOICVDNNTCDVMHIFHSGIPQSLQHLIEDSTLVKVGIGIDGDSVKLFDHYGVSIXKVED 120
OY 212 LSLDAOKIGGDKKMGKSLATETLVCKELKPRIRIKGNMEFPLSKOOLQYATDATAYAS 271
DB 121 LSLDAOKIGGDKKMGKSLATETLVCKELKPRIRIKGNMEFPLSKOOLQYATDATAYAS 180
OY 272 WHLYKVLKDLPAVSGS 288
|||||

DB 181 WHLYKVLKDLPAVSGS 197
RESULT 6
AAG24979
ID AAG24979 standard; Protein; 177 AA.
AC AAG24979;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28858.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134421.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138545.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 60.1%; Score 896; DB 21; Length 177;
Best Local Similarity 96.6%; Pred. No. 1,3e-87;
Matches 171; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 112 MOLIKVLDFTKRDSEGTAFVLDIEMRPSFKGVLPKQVATVOICVDSNVCDDVWHIFHSGT 171
DB 1 MOLIKVLDFTKRDSEGTAFVLDIEMRPSFKGVLPKQVATVOICVDSNVCDDVWHIFHSGT 60
QY 172 POSLOHLIEDSTLVKVGIGIGIDGDSVKLFHDYGVSIKDVEDLSLANQKIGCDKKGSLAST 231
DB 61 POSLOHLIEDSTLVKVGIGIGIDGDSVKLFHDYGVSIKDVEDLSLANQKIGCDKKGSLAST 120
QY 232 TETLVCKEILKPNRIRLGNMEFYPLSKOOLQYATDAYSWMILYKVLKDLPAVSGS 288

Db 121 TESTVCKELLPKRIKLNWFEYPLSKOOLQYATDAYSAMHLYKVLKDLDPDAVS DS 177

RESULT 7
AAV14517
ID AAV14517 standard; Protein; 1432 AA.

AC AAV14517;

DT 31-AUG-1999 (first entry)

DE Human WRN gene product.

KM Human; WRN; Werner's syndrome; detection; diagnosis; autosomal;
KW recessive disorder; phenotype.

OS Homo sapiens.

PN WO9724435-A1.

PD 10-JUL-1997.

PF 30-DEC-1996; 96MO-US20785.

PR 12-APR-1996; 96US-0632175.

PR 29-DEC-1995; 95US-0009409.

PR 29-DEC-1995; 95US-0580535.

PR 30-JAN-1996; 96US-0010835.

PR 30-JAN-1996; 96US-0594242.

PA (DARW-) DARWIN MOLECULAR CORP.

PA (OSHI/) OSHIMA J.

PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;

DR WPI; 1997-363671/33.

DR N-PSDB; AAX83001.

PT Isolated nucleic acid molecule encoding the WRN gene product -

PT useful for detection and treatment of Werner's syndrome, and related

PS diseases

PS Claim 10; Fig 2B; 153pp; English.

CC This sequence represents the human WRN gene product which is associated

CC with Werner's syndrome. The products can be used for the detection and

CC treatment of Werner's syndrome (WS), an autosomal recessive disorder

CC with a complex phenotype, as well as related diseases.

CC

CC

CC

CC

CC

CC

CC

CC

RESULT 8
AAW59454
ID AAW59454 standard; Protein; 1401 AA.

AC AAW59454;

DT 28-AUG-1998 (first entry)

DE Mouse WRN helicase protein.

KM Werner's syndrome; WRN; helicase; murine; diagnosis.

OS Mus sp.

PN JP10146188-A.

PD 02-JUN-1998.

PF 15-NOV-1996; 96JP-0304721.

PF 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

PR 15-NOV-1996; 96JP-0304721.

RESULT 9
AAW97841
ID AAW97841 standard; Protein; 1401 AA.

AC AAW97841;

DT 07-JUN-1999 (first entry)

DE Murine mWRN polypeptide (WRN homologue).

KM mWRN gene; WRN gene; homologue; Werner's syndrome; mouse;

OS life span; ageing.

PN

PD

PF

PF

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

XX MO906543-A1.
PN 11-FEB-1999.
XX
PF 03-AUG-1998; 98MO-US16081.
PR 04-AUG-1997; 97US-0054629.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Guarente LP, Lombard DB, Sinclair DA;
XX
DR MPI; 1999-153773/13.
DR N-PSDB; AAX24302.
XX
PT Identifying agents that inhibit replication and accumulation of DNA
PT circles - as potential agents for increasing the life span of
PT animals
XX
XX Claim 13; Page 69-72; 78pp; English.
XX
CC This is the amino acid sequence of a polypeptide encoded by mWRN
CC see AAX24302), the murine homologue of the human WRN gene associated
CC with Werner's syndrome. A claimed assay for identifying a compound
CC which extends life span comprises administering a compound to be
CC tested to a mammal with a suppressed level of mWRN, and identifying
CC slowing of at least one of the normal phenotypes of ageing in the
CC mammal. A compound identified by this method inhibits the
CC replication and/or accumulation of rDNA circles in a cell. The
CC assay can be performed using a mWRN knockout mouse (also claimed).
CC mWRN nucleic acids can be used to produce mWRN polypeptides and as
CC diagnostic probes and primers. mWRN polypeptides can be used to
CC raise antibodies (for inhibiting activity of mWRN or for determining
CC it in tissues). The assay is based on the observation that in yeast
CC cells, accumulation of ribosomal circular DNA is responsible for
CC age-related enlargement and fragmentation of the nucleolus. A
CC mutation in the yeast WRN homologue SGS1 causes premature ageing,
CC suggesting a common ageing mechanism in all eukaryotes. Insight
CC into the ageing process in model systems can provide insight into
CC ageing in humans.
XX
XX Sequence 1401 AA:
SQ
Query Match 22.7%; Score 339; DB 20; Length 1401;
Best Local Similarity 36.6%; Pred. No. 1.5e-26;
Matches 74; Conservative 42; Mismatches 72; Indels 14; Gaps 5;
QY 88 NPFAMRFGRIILSKTATEVDKRAMOLIKVLDTRKDESGIAFYGLDIEMRPSFRKGVLP 147
DB 37 NLFPLEPGSIVSYEASD---CSFLSEDISMRISDGDV--VGFMEMPPIYK---PG 86
QY 148 K---VATVQICVDNSYCDVNHIFHSGI-POSTLOHLEDSTLVKVGIGIDDSVKLFHDYG 203
DB 87 KRSRVAVIQLCVSESKCYLFHISMSVFPGGLMLENKSIKRAVGIEDDQKLLRDFD 146
QY 204 VSIKDVEDSLDNOKTIGCKWGLASLTETLVCKELKRNRIPLGNWEPYPLSKOOLY 263
DB 147 VKLESFVELDVANERKICAEIWSLNGLVKLVGLKQLDKSIRCSNWSNFPLEDDKLY 206
QY 264 AATDAYASWHLVYKVLKLDPAV 285
DB 207 AATDAYAGLIITYOKLGNLGDV 228
RESULT 10
AAY14519
ID AAY14519 standard; Protein; 1401 AA.
XX
AC AAY14519;
XX
DT 31-AUG-1999 (first entry)
XX

DE Mouse WRN gene product.
XX
XX Human; WRN; Werner's syndrome; detection; diagnosis; autosomal;
XX recessive disorder; phenotype.
XX
XX Mus musculus.
XX
XX MO9724435-A1.
XX
XX 10-JUL-1997.
XX
XX 30-DEC-1996; 96MO-US20785.
XX
XX 12-APR-1996; 96US-0632175.
XX 29-DEC-1995; 95US-0009409.
XX 29-DEC-1995; 95US-0580539.
XX 30-JAN-1996; 96US-0010835.
XX 30-JAN-1996; 96US-0594242.
XX
XX (DARN-) DARWIN MOLECULAR CORP.
XX (OSHI/) OSHIMA J.
XX
XX Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;
XX
XX MPI; 1997-363671/33.
XX
XX N-PSDB; AAX83004.
XX
XX
XX Isolated nucleic acid molecule encoding the WRN gene product
XX useful for detection and treatment of Werner's syndrome, and related
XX diseases
XX
XX Claim 10; Fig 6; 153pp; English.
XX
XX This sequence represents the mouse WRN gene product. The corresponding
XX human gene product (AAY14517) is associated with Werner's syndrome. The
XX products can be used for the detection and treatment of Werner's
XX syndrome (WS), an autosomal recessive disorder with a complex phenotype,
XX as well as related diseases.
XX
XX Sequence 1401 AA:
SQ
Query Match 22.1%; Score 330; DB 18; Length 1401;
Best Local Similarity 36.0%; Pred. No. 1.4e-25;
Matches 72; Conservative 43; Mismatches 71; Indels 14; Gaps 5;
QY 88 NPFAMRFGRIILSKTATEVDKRAMOLIKVLDTRKDESGIAFYGLDIEMRPSFRKGVLP 147
DB 37 NLFPLEPGSIVSYEASD---CSFLSEDISMRISDGDV--VGFMEMPPIYK---PG 86
QY 148 K---VATVQICVDNSYCDVNHIFHSGI-POSTLOHLEDSTLVKVGIGIDDSVKLFHDYG 203
DB 87 KRSRVAVIQLCVSESKCYLFHISMSVFPGGLMLENKSIKRAVGIEDDQKLLRDFD 146
QY 204 VSIKDVEDSLDNOKTIGCKWGLASLTETLVCKELKRNRIPLGNWEPYPLSKOOLY 263
DB 147 VKLESFVELDVANERKICAEIWSLNGLVKLVGLKQLDKSIRCSNWSNFPLEDDKLY 206
QY 264 AATDAYASWHLVYKVLKLDPAV 283
DB 207 AATDAYAGLIITYOKLGNLGD 226
RESULT 11
ABB71575
ID ABB71575 standard; Protein; 346 AA.
XX
AC ABB71575;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide seq ID NO 41517.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;

XX Environmental stress resistance; salt; heat; desert; transgenic plant.
KW Salsola komarovii.
OS WO200106006-A1.
PN 25-JAN-2001.
XX 19-JUL-2000; 2000WO-JP04862.
XX 19-JUL-1999; 99JP-0235910.
PR 24-MAR-2000; 2000JP-0085377.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA Yamada A, Ozeki Y, Satou T;
XX WPI, 2001-147355/15.
DR N-PSDB; AAF74217.
XX Screening method to obtain DNA encoding environmental stress resistance
PT factor, useful for producing transgenic plants resistant to
PT environmental stress -
XX Claim 64; Page 159-160; 167pp; Japanese.
XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins
CC AAB80608 - AAB80639, which impart environmental stress resistance. The
CC invention relates to a method for identifying DNA encoding proteins
CC imparting environmental stress resistance. The method comprises inserting
CC cDNA from a library originating in a salt-resistant organism into a host
CC cell, culturing the transformants under conditions in which the
CC untransformed host does not grow well, and selecting for viable clones.
CC The method is useful for obtaining DNA encoding environmental stress
CC resistance factors. The DNA encoding proteins conferring environmental
CC stress resistance, can be used in the production of plants resistant to
CC environmental stress, which can be cultivated in unfavourable
CC environments such as deserts, salt damaged ground, cold regions and the
CC oceans. They can be used for increasing the area of land covered by green
CC plants, and desert greening and afforestation, in order to counter the
CC effects of the increase in atmospheric carbon dioxide concentration. PCR
CC primers AAF74219 and AAF74220 are used in an example illustrating the
CC method of the invention.
SQ Sequence 244 AA:
Query Match 12.0%; Score 178.5; DB 22; Length 244;
Best Local Similarity 32.0%; Pred. No. 1.7e-10;
Matches 48; Conservative 28; Mismatches 57; Indels 17; Gaps 7;
QY 130 VGLDIEMRPSFRKGVLPKATVQICVDSNYCDVMHIFHS-GIPQSHLIEDSTLVKVG 188
DB 55 IGLGVQMKRPS-----STSAATIQSLSDKK-CLIFOLSHSPAIPTLRDLDDRYTFFG 107
QY 189 I--GIDGDSVKLFHDYGVSTIKVEDLSLANOKRIGDKKMGSLTETLV--CKELKPN 244
DB 108 VHNGRARADLLQSGHNE----LDVNNLVDLAEENGHYKLKSMEDMADVIGFC-GVHKPR 162
QY 245 RIRLGMMERPLSKOOLQVATATDAVYSWML 274
DB 163 KVMLSGMDQYCLSNDOYVACVADAVYSLRL 192
RESULT 14
ABBI1264
ID ABBI1264 standard; peptide; 127 AA.
AC ABBI1264;
XX 11-JAN-2002 (first entry)
DT Human secreted protein homologue, SEQ ID NO:1634.
DE

XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antisthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor.
XX Homo sapiens.
OS WO200157188-A2.
PN 09-AUG-2001.
PD 05-FEB-2001; 2001WO-US03800.
PF 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX (HYSB-) HYSB INC.
PA Tang YT, Liu C, Drmanac RT;
XX WPI, 2001-457740/49.
DR N-PSDB; ABA08508.
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX Claim 20; Page 161; 1963pp; English.
XX Sequences ABBI0981-ABBI1230 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g. asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g. osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g. of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides

CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

SO Sequence 127 AA;

Query Match 10.1%; Score 151; DB 22; Length 127;
Best Local Similarity 35.0%; Pred. No. 5.4e-08;
Matches 41; Conservative 20; Mismatches 48; Indels 8; Gaps 2;

OY 171 IPQSLQHLIEDSTLVKVGIGIDGDSVKLFPHDYGVSIRKVEDLSLA-----NQKIGDCK 225

DB 3 LPRLLDILADGTLTKVGVCSGSDASKLQDYGIVRGCDLRYLAMRQRNNLLCNG--- 59

OY 226 WGLASLTETLVCKELKPNRIRLGNMEEFPLSKQOLQYAATDAYASMHLYKVLKDL 282

DB 60 LSLKSLAEFTVINFPLDKSLLRCSNMDAETLTEDOVITYAARDAQISVALFLHLGYP 116

RESULT 15

AA29336
ID AAY29336 standard; Protein: 147 AA.

AC AAY29336;

DT 29-SEP-1999 (first entry)

DE Human secreted protein clone cs756_2 alternate reading frame protein.

KW Human; secreted protein; nutrition; cytokine; cell proliferation;
KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
KW tumour invasion suppressor; tumour inhibition.

OS Homo sapiens.

PN W09937674-A1.

PD 29-JUL-1999.

PF 21-JAN-1999; 99WC-US01404.

PR 20-JAN-1999; 99US-0235609.

PR 22-JAN-1998; 98US-0072134.

PA (GEMV) GENETICS INST INC.

PI Agostino MJ, Clark HF, Collins-Racie LA, Fechtel K;
PI Jacobs K, Lavallic ER, McCoy JM, Merberg D, Steininger RJ;
PI Treacy M, Wong GG;

XX WPI: 1999-458682/38.

DR N-PSDB; AAX90441.

PT New polynucleotides encoding secreted human proteins derived from,
PT e.g. fetal brain potentially used as immunostimulators.

PS Disclosure; Page 133-134; 139pp; English.

XX The present sequence represents a human secreted protein. Human secreted
CC protein polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful
CC for gene therapy.

XX
SQ Sequence 147 AA;

Query Match 10.1%; Score 151; DB 20; Length 147;
Best Local Similarity 35.0%; Pred. No. 6.8e-08;
Matches 41; Conservative 20; Mismatches 48; Indels 8; Gaps 2;

OY 171 IPQSLQHLIEDSTLVKVGIGIDGDSVKLFPHDYGVSIRKVEDLSLA-----NQKIGDCK 225

DB 23 LPRLLDILADGTLTKVGVCSGSDASKLQDYGIVRGCDLRYLAMRQRNNLLCNG--- 79

OY 226 WGLASLTETLVCKELKPNRIRLGNMEEFPLSKQOLQYAATDAYASMHLYKVLKDL 282

DB 80 LSLKSLAEFTVINFPLDKSLLRCSNMDAETLTEDOVITYAARDAQISVALFLHLGYP 136

Search completed: October 31, 2002, 13:41:54
Job Time : 38 secs

